# Proposed Decision Memo for Vagus Nerve Stimulation for Treatment of Resistant Depression (TRD) (CAG-00313R)

## Decision Summary

CMS is proposing that there is sufficient evidence to conclude that vagus nerve stimulation is not reasonable and necessary for treatment of resistant depression. Accordingly, we propose to issue a national noncoverage determination for this indication.

We are requesting public comments on this proposed determination pursuant to Section 731 of the Medicare Modernization Act. We are particularly interested in comments that include new evidence we have not reviewed here. After considering the public comments and any additional evidence we will make a final determination and issue a final decision memorandum.

Back to Top

## Proposed Decision Memo

TO: Administrative File: CAG-00313R
Vagus Nerve Stimulation for Treatment of Resistant Depression

FROM:

Steve Phurrough, MD, MPA
Director
Coverage and Analysis Group

Marcel E. Salive, MD, MPH
Director
Division of Medical and Surgical Services

Beverly Lofton, MHA

Lead Health Policy Analyst, Division of Medical and Surgical Services

Jyme Schafer, MD, MPH
Lead Medical Officer, Division of Medical and Surgical Services

SUBJECT: Proposed Coverage Decision Memorandum for Vagus Nerve Stimulation for Treatment of Resistant Depression
DATE: February 5, 2007

**I. Proposed Decision**

CMS is proposing that there is sufficient evidence to conclude that vagus nerve stimulation is not reasonable and necessary for treatment of resistant depression. Accordingly, we propose to issue a national noncoverage determination for this indication.

We are requesting public comments on this proposed determination pursuant to Section 731 of the Medicare Modernization Act. We are particularly interested in comments that include new evidence we have not reviewed here. After considering the public comments and any additional evidence we will make a final determination and issue a final decision memorandum.

**II. Background**

Types of Mental Disorders

Mental disorders are health conditions that are characterized by alterations in thinking, mood, or behavior (or some combination thereof) associated with distress and/or impaired functioning (Surgeon General's Report 1999). Depression is a mental disorder characterized by alterations in mood. "Mood disorders are recurrent, life threatening (due to the risk for suicide), and a major cause of morbidity worldwide" (Nestler et al. 2002). The symptoms of depression have been recognized as far back as ancient times, with Hippocrates referring to it as melancholia (Nestler et al. 2002). The diagnosis of depression is not based on objective diagnostic tests (such as biopsies or serum chemistries) but on a highly variable set of symptoms (Nestler et al. 2002). Nestler and others have suggested, "…depression should not be viewed as a single disease, but a heterogeneous syndrome comprised of numerous diseases of distinct causes and pathophysiologies" (Nestler et al. 2002; Thase 2006). In the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV), the criteria for a major depressive episode include five or more of the following symptoms, that have been present during the same 2-week period and represent a change from previous functioning, with at least one of the symptoms being either depressed mood or loss of interest or pleasure:

1. Depressed mood most of the day, nearly every day, as indicated either by subjective report (e.g., feels sad or empty) or observation made by others (e.g., appears tearful)
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated either by subjective account or observation made by others)
3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day
4. Insomnia or hypersonic nearly every day
5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down)[1]
6. Fatigue or loss of energy nearly every day
7. Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick)
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others)
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or specific plan for committing suicide.

A major depressive disorder (MDD) is characterized by one or more major depressive episodes (MDE). MDD is a serious condition with associated morbidity and mortality. Despite intensive research, the etiology of depressive disorders is not yet completely understood (Baghai et al. 2006). The origin of this illness is believed to be multifactorial with psychological, social and biological factors interacting to cause disturbed central nervous system function (Baghai et al. 2006).

Impact of Depression

Depression is common. In people older than 65 years, 1 in 6 suffer from depression (Wang et al. 2005). In the United States, the lifetime prevalence as determined by survey of MDD is approximately 16%, and the 12-month period prevalence of MDD is approximately 7% (Kessler et al. 2003). MDD is significantly associated with other psychiatric disorders, especially substance dependence, panic and generalized anxiety disorder, and several personality disorders, with 72% of patients with lifetime MDD meeting the criteria for at least one other DSM-IV disorder (Kessler et al. 2003; Hasin et al. 2005). Disparities in the treatment for MDD among minority groups are well known (Hasin et al. 2005). “Across the life span, the course of depression is marked by recurrent episodes of depression followed by periods of remission” (Surgeon General’s Report 1999). Patients with depression can experience spontaneous remission. The American Psychiatric Association (APA) practice guideline notes, “Untreated, the episode [MDE] typically lasts 6 months or longer. Some patients with major depressive disorder will eventually have a manic or hypomanic episode and will then be diagnosed as having bipolar disorder” (APA Guideline 2000). The natural course of untreated depression has rarely been examined (Schatzberg et al. 2000).

Mental health disorders in older adults differ from those who are younger. Most older patients with symptoms of depression do not meet the full criteria for major depression, with the suggestion that the standard criteria for depression may be more difficult to apply to older adults, or that older adults are reluctant to report such feelings (Surgeon General’s Report 1999). Depression in older adults occurs in a complex psychosocial and medical context: the prevalence of clinically significant depression in later life is estimated to be highest (about 25%) in those with chronic illness, particularly those with ischemic heart disease, stroke, cancer, chronic lung disease, arthritis, Alzheimer’s disease, and Parkinson’s disease (Surgeon General’s Report 1999). Other stressful events such as the loss of friends and loved ones also increase as one ages. Bereavement is an important and well-established risk factor for depression (Surgeon General’s Report 1999). Unfortunately, a significant number of older adults with depression are not diagnosed or treated in the primary care setting (Surgeon General’s Report 1999). Other barriers to treatment include: beliefs that depression and hopelessness are normal conditions with older age and difficulties presented by patients with cognitive deficits that make identification of depression in older adults challenging (Surgeon General’s Report 1999).

## Treatments for Depression

Emotions appear to be regulated in many areas of the brain, and there is no consensus as to the site of pathology for depression (Nestler et al. 2002). Some insight into chemical changes in the brain that accompany depression were discovered when two classes of medications were found (incidentally) to be effective in treating depression.[2] There are many effective treatments for depression, and an Agency for Healthcare Research and Quality (AHRQ)-sponsored expert panel concluded that, “depression, once identified, can almost always be treated successfully” (Agency for Healthcare Research and Quality 2000). Practice guidelines for the treatment of major depressive disorder recommend pharmacotherapy, psychotherapy, psychotherapy plus pharmacotherapy, or electroconvulsive therapy. Most of the time pharmacotherapy is the first-line treatment for MDD, though, “Choosing the agent that is most appropriate for a given patient is difficult” (Hansen et al. 2005). Pharmacologic treatment for MDD includes first-generation antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) and second-generation antidepressants. Second-generation medications include: selective serotonin reuptake inhibitors (SSRIs); selective norepinephrine reuptake inhibitors, and other drugs that selectively affect the activity of neurotransmitters. There appears to be a latency of several weeks until depressive symptoms are acceptably diminished with the current pharmacotherapies. In the geriatric population, “…prescribing guidelines … are rarely based on studies actually conducted in elderly populations and often must extrapolate from studies in younger age groups” (Wang et al. 2005). Taylor states, “Studies in younger populations may not generalize to the older population as depression in the elderly differs from depression in younger individuals” (Taylor et al. 2004). Though effective treatments exist, the AHRQ-sponsored expert panel also notes, “…appropriate treatment continues to be a pressing issue” (Agency for Healthcare Research and Quality, 2000).

“Patients receiving antidepressant monotherapy may be partially or totally resistant to treatment in 10 to 30 percent of cases” (Cadieux 1998). There are several hypotheses for this therapy resistance, including: occult medical conditions causing depression, substance abuse interfering with treatment, noncompliance, abnormal metabolism, psychosocial factors, and other psychiatric comorbidities (Fava 2003; Fleck et al. 2005). Another common cause of treatment failure is prescribing antidepressant medication in dosages that are too low and for inadequate lengths of time (Cadieux 1998). Numerous studies have documented relatively low rates of adequate prescribing in various treatment settings (Cadieux 1998). For instance, a managed care setting documented adequate antidepressant therapy in only 11% of patients (Nemeroff 1996). Even in patients who have been hospitalized for major depression, Oquendo noted, “Antidepressant treatment of depressed patients is strikingly inadequate, even in suicide attempters, known to be at higher risk for suicidal acts” (Oquendo et al. 2002). Strategies after failing a standard first line treatment are drug substitution, combination strategies (the addition to a second agent), augmentation strategies (such as thyroid hormone, benzodiazepines, estrogen, dexamethasone, or lithium), or electroconvulsive therapy (ECT). Other novel treatments are under investigation (Baghai et al. 2006). Unfortunately, “There is little evidence to guide the management of depression that has not responded to a course of antidepressants” (Stimpson et al. 2002).

The definitions of treatment resistance, treatment response, and remission are variable. Treatment resistant depression (TRD) is not defined in the Diagnostic and Statistical Manual of Mental Disorders-Fourth Version (DSM-IV). Rush (2003) proposed that, “Difficult to treat depression includes depression that inherently does not respond satisfactorily to one or more treatments that are optimally delivered (treatment-resistant depression [TRD]) and also depression treated under circumstances precluding the optimal delivery of potentially effective treatments. Such circumstances include the use of subtherapeutic doses; nonadherence; intolerable side effects that prevent an adequate dose or duration of treatment; and concurrent Axis I, II, or III conditions that reduce the likelihood of remission for adherence, pharmacokinetic, or pharmacodynamic reasons) [sic]”(Rush et al. 2003). Additionally, Thase and Rush (1997) propose a model of staging for levels of resistance of TRD, though Fava (2003) states about this model, “…its predictive value with respect to treatment outcomes has not yet been assessed systematically” (Fava 2003; Thase et al. 1997). In a systematic review by Stimpson on interventions for treatment-refractory depression, these two points are included in their conclusions (Stimpson et al. 2002):

- “In the absence of good evidence, clinicians will have to rely upon their own clinical judgment in deciding upon treatment.”
- “The main conclusion is that further research is required as the findings are not strong enough to support any clinical guidance.”

Stimulation of the brain with electricity in a living person was first documented in 1874 (Gildenberg 2004). The first reported use of VNS was in 1883 by a neurologist, James L. Corning (Groves & Brown 2005). In the 1880's he performed transcutaneous stimulation over the area of the vagus nerve and observed a decrease in seizures (Gildenberg 2004). The vagus nerve, the tenth cranial nerve, has parasympathetic outflow that regulates the autonomic (involuntary) functions of heart rate and gastric acid secretion, and also includes the primary functions of sensation from the pharynx, muscles of the vocal cords, and swallowing. It is a nerve that carries both sensory and motor information to the brain. Importantly, the vagus nerve has influence over widespread brain areas (Groves et al. 2005). In 1997, a VNS device was approved by the FDA for the treatment of seizures in patients with refractory epilepsy. In a study of eleven epilepsy patients, improvement in mood was noted which lead to a suggestion of the use of VNS for depression and further studies (Elger et al. 2000). The VNS device consists of three parts: 1) a programmable pulse generator which is implanted subcutaneously in the left chest wall 2) two electrodes that are wrapped around the vagus nerve and attached to the pulse generator and 3) a programming wand for the purpose of noninvasive device programming, device diagnostics, and data retrieval. Vagus nerve stimulation is being investigated as a treatment for the cognitive impairment associated with Alzheimer's disease, anxiety, obesity, autism, migraines, involuntary movement disorders, and obsessive-compulsive disorder (Aetna Clinical Policy Bulletins 2006; Groves & Brown 2005). The precise mechanism of action of VNS remains unknown (Salzman 2006).

**III. History of Medicare Coverage**

CMS currently provides coverage for VNS for patients with medically refractory partial onset seizures, for whom surgery is not recommended or for whom surgery has failed. VNS is not covered for patients with other types of seizure disorders which are medically refractory and for whom surgery is not recommended or for whom surgery has failed (§160.18 of the Medicare National Coverage Determination Manual).

Medicare does not currently have a national coverage determination (NCD) on VNS for treatment of resistant depression (TRD). In absence of an NCD, coverage is determined by local Medicare contractors.

**Current Request**

On July 26, 2006, CMS received a formal request for reconsideration from Cyberonics, Inc. The company proposed that CMS revise its current NCD to include coverage of VNS for TRD for patients who have been either (1) previously treated with or refused treatment with electroconvulsive therapy (ECT), or (2) have been previously hospitalized for depression. The specific indication requested for coverage is for the adjunctive long-term treatment of chronic or recurrent depression for patients over the age of 18 who are experiencing a major depressive episode and have not had an adequate response to four or more adequate depression treatments.

Cyberonics, Inc. also requests the following to be considered as contraindications:

- The VNS Therapy System cannot be used in patients after a bilateral or left cervical vagotomy.

- Do not use short wave diathermy, microwave diathermy or therapeutic ultrasound diathermy on patients implanted with a VNS Therapy System. Diagnostic ultrasound is not included in this contraindication.

**Benefit Category**

For an item or service to be covered by the Medicare program, it must meet one of the statutorily defined benefit categories outlined in the Social Security Act. Vagus Nerve Stimulation, at a minimum, falls under the benefit categories set forth in sections §1861(s) (6) (durable medical equipment), 1861(s) (q) (physicians' services), and 1861(s) (2) (B), (hospital services "incident to" physicians' services rendered to outpatients). This may not be an exhaustive list of all applicable Medicare benefit categories for this item or service.

**IV. Timeline of Recent Activities**

| | |
|---|---|
| July 26, 2006 | CMS received a formal request for reconsideration from Cyberonics, Inc. to include coverage of VNS for treatment of TRD. |
| August 7, 2006 | CMS formally opened a national coverage determination (NCD) as reconsideration to be made on VNS. |

| | |
|---|---|
| | The initial public comment period opened. |
| September 6, 2006 | The initial public comment period closed. |
| October 30, 2006 | Cyberonics, Inc. meeting with CMS. |

**V. FDA Status**

FDA approval for the VNS Therapy System was received on July 15, 2005. This device is indicated for the adjunctive long-term treatment of chronic or recurrent depression for patients eighteen years of age or older who are experiencing a major depressive episode and have not had an adequate response to four or more adequate antidepressant treatments.

**VI. General Methodological Principles**

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary.

Methodological principles of study design that are used to assess the literature on a therapeutic or diagnostic item or service for specific conditions can be found in Appendix A. In general, features of clinical studies that improve quality and decrease bias include the selection of a clinically relevant cohort, the consistent use of a single good reference standard, and the blinding of readers of the index test, and reference test results.

Public comment sometimes cites the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

**VII. Evidence**

**A. Introduction:**

The evidence provided by the sponsor included a randomized controlled trial for FDA PMA approval (D02), a case series study (D01), a company sponsored observational study (D02 observational), and a company sponsored trial of standard treatment for depression (D04) that was used as a comparison study for D02. The sponsor provided information in addition to the previously mentioned evidence: a booklet of sponsor commentary, a study of VNS in rats, abstracts, physiology studies, economic information including a cost analysis, investigator biographies, reviews of TRD, reviews of VNS, overview of MDD, STAR*D publications, an ECT trial, posters, and letters to insurance companies.

Assessment of Outcome in Depression

The use of outcomes measures attempt to follow what happens to a patient over time to quantify what is happening during the course of treatment. This is used for comparison purposes and to monitor a patient's progress and treatment. The outcomes of interest for treatment with the VNS device for TRD are improvement in depression and implant-related adverse events. While we have a concept of what depression is, we cannot measure changes in depression directly. We can examine depressive symptoms, social and work functioning, quality of life, morbidity such as hospitalization, and mortality. Standardized outcome instruments commonly used involve a measurement of depressive symptoms in which patients' subjective experiences are translated into a numeric rating scale. Some of these symptom-only instruments include: the Hamilton Depression Rating scale (HDRS or HAM-D or HRSD), the Montgomery-Asberg Depression Rating Scale (MADRS), the Inventory of Depression Symptomatology Scale (IDS) (either clinician administered – CR, or self-administered –SR), and the Beck Depression Inventory (BDI). Depressive symptoms have been used to create items that make up scales, where the assumption is that the presence and absence of these symptoms and the patterns in which they occur define the illness—depression—and the outcome is the scale score from item answers. Mathematical operations on these numbers are assumed to reflect true patient changes (Bech 2006). A number of scales (hundreds) have been developed and are reliant on these assumptions (Veterans Administration 2004). In 1979 Montgomery et al noted, "The large number of rating scales available to clinical investigators is a problem in psychiatric research (Pichot, 1972) and the comparability between scales is rarely known" (Montgomery et al. 1979).

The purpose of using scales to measure depression can vary. Such measures may be used for screening or diagnosis, or as a tool for outcome assessment. To measure change brought about by treatment (outcome), the ability to detect small but clinically meaningful differences in severity is important (Nelson et al. 2006). Analysis of the scale as an outcome includes asking two questions: is the score meaningful and is the scale meaningful? Using metrics that are both reliable and valid is important. Reliability and validity determination is both an art and a science. Validity refers to the degree which a test measures what it intends to measure. Reliability examines the consistency between two measures that evaluate the same thing, and is the ratio of the true variance to the total variance. There are several methods to assess reliability: examining internal consistency (how well do scale items measure a single characteristic); retest reliability (assesses to what degree multiple administrations of the scale produce the same results); and interrater reliability (the degree to which various raters produce the same result)(Bagby et al. 2004). Reliability is a group-specific statistic, so if a narrowly defined population has small variance in the true score, the metric will be less reliable in a different population. Bagby et al have analyzed the psychometric properties of the most commonly used measure of depression, the original 17 item Hamilton Depression Rating Scale (Bagby et al. 2004).

Though it has been in use for 40 years, some suggest that the original 17-item version may be problematic from a reliability and validity standpoint (Bagby et al. 2004; Licht et al. 2005). Rehm et al report that this scale was developed not from a statistical or empirical process, but from logic (Rehm et al. 1985). Bagby et al state, “Finally, the Hamilton depression scale is measuring a conception of depression that is now several decades old and that is, at best, only partly related to the operationalization of depression in DSM-IV” and additionally, “In conclusion, we have been struck with the marked contrast between the effort and scientific sophistication involved in designing new antidepressants and the continued reliance on antiquated concepts and methods for assessing change in the severity of the depression that these very medications are intended to affect.” Some authors have noted that patients with equivalent total scores may have very different symptoms and thus, different meaning (Bech 2006; Nelson et al. 2006). The Hamilton items 18-21 are: diurnal variation, depersonalization, paranoid symptoms, and obsessive and compulsive symptoms. The 24-item scale adds hopelessness, worthlessness and helplessness (Nelson et al. 2006). In the review by Nelson, he stated of the 24-items, “These three items, however, have received much less attention in the literature and at this point do not have the empirical support that the other core symptoms have.” The MADRS (1979), the second most popular scale in antidepressant studies, was developed to be more sensitive to change, however, this scale was developed during the tricyclic antidepressant time (reflecting the symptom changes associated with this treatment) and many in this sample were inpatients (Nelson et al. 2006). Nevertheless, some suggest that the MADRS is superior to the HRSD 17 item for clinical trial outcomes (Carmody et al. 2006). The less frequently used IDS 28-item was published by Rush et al. in 1986 (Veterans Administration 2004; Nelson et al. 2006). In summation, it is not clear which items and which scales are most sensitive at measuring changes during a patient’s treatment for depression (Nelson et al. 2006). The Young Mania Rating Scale (YMRS) is the most frequently used scale for assessing mania severity in patients already diagnosed with mania.

Instruments that are not depression-specific can also be used. The Global Assessment of Function (GAF) is a generic scale that evaluates both symptoms and functioning. It is not used much in the depression literature, but was a high ranking global instrument in Veterans Administration Technology Assessment Program (VATAP) report examining outcomes measurement in major depression to use in measuring the quality of treatment in Veterans Health Administration (VHA) mental health services. The SF-36 is a generic measure of perceived health status that also measures symptoms and function. The Clinical Global Impression (CGI) scale refers to the global impression of the patient, a single item response (normal to extremely ill). The VATAP has created a measures evaluation matrix of fifteen instruments that met certain criteria for depression treatment outcomes measurement

(http://www.va.gov/vatap/pubs/Depressionfinal3-05.pdf). CMS generally accords more weight to outcomes with validated measures of patient functioning (social and work), quality of life, morbidity (such as hospitalization), and mortality.

There is a lack of empirical evidence for endpoints in clinical studies of depression (Rush et al. 2006). Concepts of response, remission, recovery, relapse, and recurrence do not have standardized, empirical definitions (Rush et al. 2006). The recommended end point in the treatment of depression has become remission, but remission has been defined in a variety of ways, including: a score of 7 or less on the HAM-D 17 item, minimal or no symptoms of depression, failure to meet the DSM-IV diagnostic criteria for MDD, or return of normal function (both social and occupational) (Zajecka 2003). Rush et al 2006 note, "Thus, use of different operational definitions of remission can lead to radically different descriptions of the course of illness, including both the number and duration of MDEs" (Rush et al. 2006). Rush et al. recommend that response criteria be met for 3 consecutive weeks "to take into account error in the assessment of symptomatology and unstable symptomatic fluctuations," though this recommendation is not empirical (Rush et al. 2006). The association of symptom reduction with functional improvement is not well defined (Rush et al. 2006). Rush et al. recommend that remission refers only to the symptoms noted in DSM-IV, and that 3 consecutive weeks pass, during which each week was characterized by the absence of depressive symptoms (Rush et al. 2006). Symptoms can recur but may be insufficient in number, duration, or intensity to qualify for a relapse or recurrence; again, few studies have empirically evaluated these concepts (Rush et al. 2006). Rush et al. state, "Remission typically follows response by at least several weeks," and recommended a 12-20 weeks trial duration, though in prolonged trials the chances of spontaneous remission increase, and some patients who remit will relapse (Rush et al. 2006). In general, more weight is given to conclusions of studies that have a scientifically derived endpoint of remission.

Well designed clinical trials are important for accurate outcome interpretation. Well constructed randomization protects against bias and inclusion of an appropriate comparator facilitates study interpretation. The placebo effect is a substantial, common consideration in trials of antidepressants. About one-half of randomized, double-blind placebo controlled antidepressant trials fail to show statistical superiority of widely used antidepressant in comparison to placebo (Khan et al. 2002). In 1999, a National Depressive and Manic-Depressive Association consensus statement on the use of placebo in clinical trials of mood disorders concluded that placebo has a role in mood disorder studies and new drugs found to be equivalent to standard treatment are not evidence of efficacy unless the new drug is significantly more effective than placebo (Kupfer et al. 2002). The placebo effect appears to be a complex mental activity, having different mechanisms in different conditions, meaning there is not a single effect but many (Benedetti et al. 2005). In patients being treated with antidepressants, placebo response patients have been observed to have similar PET scan changes (Mayberg et al. 2002). Benedetti et al provide these comments: "The study of the placebo effect reflects a current neuroscientific thought that has as its central tenet the idea the 'subjective' constructs such as expectation and value have identifiable physiological bases, and that these bases are powerful modulators of basic perceptual, motor, and internal homeostatic processes,"; " …the existence of placebo effects suggests that we must broaden our conception of the limits of endogenous human capability" (Benedetti et al. 2005). In the case of research in the area of depression, more weight will normally be accorded to studies that are designed to guard against the placebo effect.

Adverse events are important medical outcomes. Patients need this information to make well-informed choices. For instance, invasive procedures such as the implantation of the VNS device in the carotid artery sheath could include events such as infection and tissue scarring. Serious injuries such as vocal cord paralysis, sleep apnea, shortness of breath, syncope, cardiac arrhythmias, and difficulty swallowing could be examples of potential adverse outcomes. Studies that provide an inclusive examination and explanation of adverse medical events are generally given more weight.

**B. Discussion of evidence reviewed**

**1. Question:**

The development of an assessment in support of Medicare coverage decisions is based on the same general question for almost all requests: “Is the evidence sufficient to conclude that the application of the technology under study will improve health outcomes for Medicare patients?” For this NCD, the question of interest is:

Is the evidence sufficient to conclude that, in the Medicare population, vagus nerve stimulation will improve health benefits for individuals with treatment resistant depression?

**2. External technology assessments**

CMS did not commission an external technology assessment (TA); however, external assessments were identified on the topic of vagus nerve stimulation for treatment resistant depression**.**

Blue Cross Blue Shield Technology Evaluation Center (TEC)

August 2005 the TEC published a TA titled, Vagus Nerve Stimulation for Treatment-Resistant Depression. Vagus nerve stimulation for treatment-resistant depression met only one of five of the TEC criteria. TEC determined, "…VNS therapy for the indication of treatment-resistant depression does not meet the TEC criteria." The following TEC criteria were not met: 1) the scientific evidence must permit conclusions concerning the effect of the technology on health outcomes; 2) the technology must improve the net health outcome; 3) the technology must be as beneficial as any established alternatives; and, 4) the improvement must be attainable outside the investigational setting.

These were some of the TEC concerns:

"Overall, the evidence supporting efficacy of VNS is not strong. The single randomized clinical trial did not show statistically significant results in favor of VNS for the primary outcome. Treatment response in the randomized clinical trial was much lower than had been observed in case series studies, raising concerns about placebo effects and observer bias. The non-randomized observational study had numerous methodological problems. Alternative analyses showed diminished or no efficacy of VNS therapy. Although the FDA voted to approve VNS therapy, a poll of committee members showed that approval was based on the safety of VNS therapy rather than strong evidence of efficacy."

"Patient selection was a concern for all studies. VNS is intended for treatment-refractory depression, but the entry criteria of failure of 2 drugs and a 6-week trial of therapy may not be a strict enough definition of treatment resistance. Treatment-refractory depression should be defined by thorough state-of-the-art psychiatric evaluation and management before an invasive surgical procedure of limited efficacy is performed."

In summary, the TEC report stated the following.

"…The available evidence is not sufficient to permit conclusions of the effect of VNS therapy on health outcomes."

The TEC updated its assessment in August 2006, with no change in TEC criteria determination. The TEC again concluded:

“Since the last TEC Assessment, there have been no studies reporting clinical outcomes on any new or different patients. Data from the case series and clinical trials have been reanalyzed to show what proportions of patients who respond at one time are still responders at a subsequent time point. However, this information by itself does not provide evidence of the efficacy of VNS beyond that provided by the original observational comparison of VNS versus treatment as usual.”

California Technology Assessment Forum (CTAF)

February 2006 the CTAF published a TA titled, Vagus Nerve Stimulation for Treatment-Resistant Depression. Vagus nerve stimulation for treatment-resistant depression met two of the five CTAF criteria. The two criteria met were: 1) the technology must have the appropriate regulatory approval; and 2) the scientific evidence must permit conclusions concerning the effectiveness of the technology regarding health outcomes. The following three CTAF criteria were not met: 1) the technology must improve the net health outcome; 2) the technology must be as beneficial as any established alternatives; and, 3) the improvement must be attainable outside the investigational settings. Based on the criteria, the CTAF panel approved the recommendation that “…VNS does not meet Technology Assessment Criterion 3, 4 or 5 for effectiveness and improvement in health outcomes for TRD.”

In summary the TEC report stated the following.

“… given the well-designed negative RCT and the fact that this is a single group of patients in an observational trial, it is early to conclude that the new technology of VNS improves the net health outcomes as much as or more than the established alternative of TAU with medications and/or electroconvulsive therapy.”

**3. Internal technology assessments**

CMS performed a literature search utilizing PubMed evaluating the use of VNS for the treatment of depression and a review of end references. The literature search was limited to the English language and specific to the human population. Public access information from the FDA website was also used. Technical assessments using VNS for depression were searched for using Google.

Evidence for VNS for the treatment of TRD comes from the randomized controlled trial for FDA PMA approval (D02), a case series study (D01), and a company sponsored observational study (D02 observational). The evidence table is located in Appendix B.

**Evidence Summary**

**VNS Pivotal Study**

D02 was a randomized multi-center placebo-sham trial with the objective of evaluating patients at 12 weeks post-implantation and during a 12 month follow-up period. The control-sham group received a functional VNS device, which was to be turned on after 12 weeks. Rush et al. state, “The sample size was powered to detect a difference in the HRSD 24 response rate of approximately 17%.” The 12 week trial (acute phase) was extended (long-term phase), with the control group’s devices activated.

**D02 Randomized Controlled Trial**

This 3 month double-blind trial randomized 235 outpatients with major depressive disorder (n = 210) or bipolar disorder, depressed phase (n = 25) to either active VNS treatment or inactive VNS treatment (sham control) at 21 sites. A third party assigned sequential numbers to all subjects and randomized the subjects 1:1. The device programmer had the randomization assignment for each participant so the active treatment devices could be activated (the programmer was not involved in care or clinical assessment, however, the programmer did collect information on all adverse events). Control patients with inactive devices had follow-up visits with the intent of device adjustment, and investigators were blinded to treatment. Inclusion and exclusion criteria based on Rush et al. 2005, George et al. 2005, and the FDA Clinical Memorandum are presented in Appendix C.

Patients were randomized to either active VNS or inactive VNS. Patients had a 2 week recovery after implantation, followed by 2 weeks of electrical parameter adjustment and then 8 weeks of fixed electrical stimulation. For the adjustment of the electrical parameters, "…output current (mA setting) was increased progressively to the maximal level that could be comfortably tolerated by the participant" (Rush et al. 2005a). Initial electrical treatment parameters (frequency in hertz, pulse width in microseconds, on-off cycle in seconds and minutes, and output current in milliamps) were identical to those used for patients with epilepsy. Medication changes or ECT were not allowed other than the addition of the antidepressant trazodone (up to 300 mg/day). The primary efficacy endpoint was the proportion of subjects who had $\geq$ 50% decrease in the HAM-D 24 at visit 9 (12 weeks after implantation, 10 weeks of VNS therapy) as compared to the baseline value (FDA Clinical Memorandum). Protocol violators ("if they did not complete the acute phase, discontinued for reasons other than treatment-related adverse events or lack of efficacy, if implanted and had concomitant anti-depressant medication adjustments for at least 7 days during the acute phase, or received ECT during the acute phase") were not considered in the efficacy analysis (FDA Clinical Memorandum). After the two week adjustment period, patients were seen weekly for two weeks then every other week over the following 6 weeks. Rush et al. states, "Efficacy and safety data were gathered at the two baseline visits and at post-implantation weeks 1 and 2 (recovery period), weeks 3 and 4 (stimulation adjustment period), and weeks 5, 6, 8, 10, and 12 (fixed-dose stimulation period)." Response measures were differentially evaluated (FDA Clinical Memorandum). HAM-D 28, MADRS, CGI, IDS-SR, and YMRS had 2 assessments during baseline (FDA Clinical Memorandum). HAM-D 28, MADRS, CGI, IDS-SR had 1 assessment during recovery and YMRS had 2 assessments during recovery (FDA Clinical Memorandum). During the remaining 10 weeks of the trial the measures were collected in the following manner: HAM-D 28 and MADRS had 4 assessments; CGI, 1 assessment at acute phase exit; IDS-SR and YMRS, 5 assessments. SF-36 was collected at baseline and acute phase end (FDA Clinical Memorandum). Rush et al. states, "Although the 28-item HRSD was administered to participants, the total of the first 24 questions was used to define the HRSD 24 total score." Although multiple secondary outcomes were collected, no adjustments were made for multiple comparisons. Safety was assessed by evaluation of adverse events, serious adverse events (death, life-threatening event, in-patient hospitalization or prolonged existing hospitalization, persistent or significant disability/capacity), and physical and neurological examinations (FDA Clinical Memorandum ).

**Table 1:** D02 Acute Phase Enrollment(FDA Clinical Memorandum)

| Tracking Point | Subject Number |
|---|---|
| Target Enrollment | 275 |
| Actual Enrollment | 266 |
| | |

| Tracking Point | Subject Number |
|---|---|
| Discontinued pre-implant | 31 |
| Implanted | 235 |
| Discontinued Acute Phase* | 13 |
| End of Acute Phase (evaluable subjects) | 222 |
| Randomized subjects | 222 |
| Treatment Group | 112 |
| Control Group | 110 |

* 13 patients discontinued the 12 week acute phase: 4 did not meet visit 2 continuation criteria; 9 were protocol violators.

Demographics revealed a mean age of 46.3 (N = 205), 74/205 (36%) male, 198/205 (97%) Caucasian (FDA Clinical Memorandum). Rush et al. stated, “Demographic data are reported on the 222 evaluable participants, and safety findings are reported on the total 235 implanted participants,” reporting 96% Caucasian, 63% female, mean age 46.5 years (SD 9.0), median 47.0 years (range 24-72). Concomitant treatments were supposed to remain stable, however 9 subjects (four treatment subjects and five control subjects) had changes in antidepressant, atypical antipsychotic, or anticonvulsant medications, and were therefore protocol violators (FDA Clinical Memorandum). Additionally, 3 subjects had increases in medication (FDA Clinical Memorandum). No ECT treatments were given.

After 3 months, 15% (17/111) of patients in the active VNS group met the response criteria of a 50% reduction in HRSD-24, whereas 10% (11/110) met this criteria from the placebo-sham group (p = 0.238) (FDA Clinical Memorandum). A last observation carried forward (LOCF) analysis of responders also did not reveal statistical significant difference (FDA Clinical Memorandum). Of the 21 sites, Rush et al. stated, “Response rates were generally similar across sites, although some variation was seen (seven sites had < 10% response rate, four sites had $\geq$ 25% response rate).” Of the secondary measures (IDS-SR, CGI, MADRS, SF-36) only IDS-SR had a statistically significant difference in outcome, in favor of active VNS treatment (19/109 versus 8/106, p = 0.032) (FDA Clinical Memorandum). Rush et al. also found LOCF outcomes for IDS-SR response rates to be statistically significant (treatment group 17.0%, n = 112, control group 7.3%, n = 110; p = 0.032, reporting as a footnote that one patient in the control group did not have an IDS-SR assessment completed during the study), but no statistical difference for IDS-SR % improvement from baseline (p = 0.158). It is noted that, “An exploratory analysis of this acute study found no relationship between output current and the percentage of change in the HRSD 24” (Rush et al. 2005a).

Adverse events were categorized based on implantation related, stimulation related, serious adverse events, hypomanic/manic reactions, suicidal ideation, and death (FDA Clinical Memorandum). Implantation related adverse events reported at a $\geq$ 5% incidence among all implanted patients (N = 235) were incision pain, voice alteration, incision site reaction, device site pain, device site reaction, pharygitis, dysphagia, hypesthesia, dyspnea, nausea, headache, neck pain, pain, paresthesia, and increased cough. (FDA Clinical Memorandum). Stimulation related adverse events reported at a $\geq$ 5% incidence among treatment patients (N=119) were asthenia, back pain, chest pain, device site pain, device site reaction, headache, incision pain, neck pain, pain, viral infection, would infection, palpitation, constipation, diarrhea, dyspepsia, dysphagia, nausea, vomiting, depression, dizziness, hypesthesia, insomnia, paresthesia, cough increase, dyspnea, laryngismus, pharyngitis, rhinitis, voice alteration, and incision site reaction (FDA Clinical Memorandum).

Twenty-seven patients reported 39 serious adverse events (events that required hospitalization or prolonged hospitalization, resulted in death, were considered life threatening that resulted in a persistent or significant disability or incapacity, or other) (FDA Clinical Memorandum). Rush et al state, “Of 30 total serious adverse events (SAEs) involving 27 participants, 16 SAEs occurred in the active VNS group and 14 in the sham group. This total included 12 SAEs involving 11 participants of worsened depression that required hospitalization (seven participants in sham, four participants in active VNS, and one participant who had not yet received stimulation, but who was assigned to the active VNS group).” Thirty of these events occurred after implantation, with the most common event being worsening depression (N=12, 5 in the treatment group and 7 in the sham group), and included suicide, asystole, bradycardia, confusion, thinking abnormal, aspiration pneumonia, pneumonia, and renal failure (FDA Clinical Memorandum). Three subjects had adverse events of manic reaction in this acute phase (FDA Clinical Memorandum). “Two participants in the active VNS group (one of whom had a diagnosis of bipolar I disorder at baseline) met the threshold of significant hypomania, a score $\geq$15 on the YMRS, which was validated by DSM-IV criteria” (Rush et al. 2005). Suicidal ideation was evaluated by an increase of HAM-D item 3 of at least 2 points: 3/116 of the sham group and 2/119 of the treatment group met this criterion (FDA Clinical Memorandum). One death occurred before implantation (esophageal cancer) and one death from suicide occurred during the acute phase (treatment group). Three subjects left the study because of adverse events (Rush et al. 2005).

**D02 Observational Study**

Subjects at the exit of the 3 month (visit 9) acute phase study entered into the long-term phase study (FDA Clinical Memorandum). The FDA Clinical Memorandum states that, “the purpose of the long-term analysis is to examine adverse effects that occur after long-term exposure to VNS therapy.” Patients randomized to sham therapy in the acute phase were included only if the HRSD score was $\geq$ 18 at the two last assessments of the acute phase trial (average of the two assessments). Rush further stated, “they could elect to receive active VNS for humanitarian reasons.” Patients with sham devices who met this criterion had their devices activated at this time. While the baseline for the active treatment group remained the averaged ratings before implantation, the sham VNS group (the device being activated at entry into this phase of the study) had their baseline changed to the average of the 8 and 10 week rating (Rush et al. 2005b). Medication changes and ECT were allowable. Device voltage adjustments and medication adjustments were allowed throughout this time period. Safety was assessed similar to the randomized acute phase trial (FDA Clinical Memorandum). In contrast with the randomized acute phase trial, protocol violators could be included in the efficacy analyses (FDA Clinical Memorandum).

**Table 2:** D02 Enrollment (acute and long-term phase) (FDA Clinical Memorandum)

| Tracking Point | Subject Number |
|---|---|
| Target Enrollment | 275 |

| Tracking Point | Subject Number |
|---|---|
| Actual Enrollment | 266 |
| Discontinued pre-implant | 31 |
| Implanted | 235 |
| Acute Phase | 235 |
| Discontinued Acute Phase | 2 |
| End of Acute Phase | 233 |
| Long-term phase* | 233 |
| | |

| Tracking Point | Subject Number |
|---|---|
| Not evaluable** | 28 |
| Evaluable Subjects | 205 |
| Not 12 Month completers*** | 28 |
| 12 month subjects who completed trial | 177 |

* Two subjects did not meet acute phase continuation criteria because they only had continuation visits, thus 231 patients could be considered as intent-to-treat (ITT) subjects.
** Twenty-eight subjects were not evaluable. Four subjects (original treatment group) had no assessment data (no HAM-D scores post-acute phase) collected at any long-term visit. Three patients (original treatment group) did not meet acute phase continuation criteria. Twenty-one patients (sham-placebo group) had a 12 week exit score of HRSD-24 $\leq$ 18, so did not meet the criteria to continue in the long-term phase.
*** Twenty-eight subjects were not considered 12 month completers. Seventeen subjects discontinued prior to one year, 6 did not have > 80% stimulation, and 5 did not have 11 or 12 month assessments.

Rush et al commented: “two subjects discontinued acute phase; one was not included due to suicide and the other because of device explanation due to infection, both occurring in the acute phase” (Rush et al. 2005a). Twenty eight patients were not evaluable: 3 had HRSD 24 scores < 18 after implantation; four participants in the initial active VNS group had no HRSD 24 scores after acute phase exit and 21 initial sham participants did not average $\geq$18 on the HRSD 24 at 8 and 10 weeks of sham treatment. Twenty-eight were not 12 month completers; 17 discontinued participation before 1 year (4 due to adverse events; 7 due to lack of efficacy, six because of other participant decisions), 6 did not have stimulation > 80% of the time, and 5 did not have 11 or 12 month assessments. The authors defined a group called the observed sample, where only participants with data available for each measurement at each time point were included, which is the reason they give as to why the number of people evaluated for different outcomes vary.

The demographics of the 205 evaluable were similar to the original 222 participants, including the revised baseline HRSD 24, MADRS, and IDS-SR 30 ratings.

Changes in concomitant medications and ECT, as well as other therapies, were allowed during this phase. VNS was monitored by the study investigator, while medications and other treatment, such as ECT or psychotherapy, were managed by the subject's regular health care provider or could be managed completely by the investigator. Fourteen subjects received ECT treatment (FDA Clinical Memorandum). By twelve months, median output current was 1.0 mA (range 0.00 to 2.25 mA).

Subjects were followed differentially. Acute phase treatment group subjects had monthly visits for 12 months. Acute phase sham treatment (delayed treatment) subjects had weekly visits for 4 weeks and then every other week for 10 weeks of VNS, then monthly visits until 12 months post implantation (FDA Clinical Memorandum). "During the long-term phase, all subjects had follow-up evaluations at months 6, 9, and 12, after implantation and stimulation adjustments were permitted; acute responders had additional follow-up assessments" (FDA Clinical Memorandum). Measures were evaluated differentially (FDA Clinical Memorandum). HAM-D, MADRS, and IDS-SR were assessed monthly up to 12 months; CGI was assessed monthly from 6 to 12 months; and YMRS and SF-36 were assessed at 6, 9, and 12 months. Though clinical raters knew the patients were on VNS therapy, they were masked to device settings and medications.

Rush et al stated, "The a priori specified primary outcome is a repeated measures analysis of the HRSD 24 total score, which estimated the average monthly change in HRSD 24 over 12 months of stimulation." The model was adjusted for baseline HRSD 24, acute study treatment group (those randomized in the acute phase), and pooled site. Unequally spaced visits were dealt with by statistical modeling. Response was defined as in the earlier phase of this study, a reduction of 50% of more in the score compared with baseline for the HRSD-24, IDS-SR30, or MADRS, or a CGI-I of 1 or 2 (much or very much improved). Remission was defined as a score $\leq$9 on the HRSD 24, $\leq$14 for the IDS-SR, or $\leq$10 on the MADRS. Treatment failures were defined as participants who exited because of VNS therapy-related adverse events or lack of efficacy, met suicide exclusion criteria, attempted suicide that results in significant (>3 days) hospitalization, or developed mania or four or more periods of rapid cycling. Additionally Rush noted, "Participants who lacked scores for an evaluation (e.g., MADRS) at 3 months could not be included in this analysis, thus accounting for the number of participants being slightly less than 205." The investigators defined what they refer to as the durability of benefit ("sustained response") as a $\geq$50% reduction in baseline HRSD 24 symptoms at least once during the last quarter (months 9, 10, 11, or 12) and achieving at least a 40% reduction from baseline on at least two other of the HRSD 24 assessments in the quarter. No statistical adjustment was made for multiple comparisons.

For the repeated measures analysis of the HRSD 24 total score (the primary analysis), 205 patients provided data at 3 months, 197 at 6 months, 186 at 9 months, and 181 at 12 months (Rush et al. 2005b). The authors stated that on average, the HRSD 24 score improved 0.45 (SE = 0.05) points per month. The group mean scores the HRSD 24, IDS-SR30 and MADRS reveal statistically significant reductions when baseline was compared with the available 12 month ratings (HRSD 24: baseline 28.0 ± 5.7 (n = 205), 12 months observed 19.6 ± 9.7 (n = 180), 12 months LOCF 20.6 ± 9.9 (n = 205); IDS-SR30: baseline 42.9 ±10.0 (n=204), 12 months observed 32.6 ± 15.3 (n = 180), 12 months LOCF 33.6 ± 15.4 (n = 204); MADRS: baseline 30.8 ± 6.9 (n = 205), 12 months observed 21.2 ± 11.5 (n = 181), 12 months LOCF 22.2 ± 11.7 (n=205). The authors stated, “We conducted several appraisals of the clinical importance of this symptomatic improvement.”

- For HRSD 24, 27% (55/202) of subjects met the response criteria at exit (LOCF analysis), and 16% (32/202) met the remission definition. For the observed population, 30% (54/181) met the response definition after 12 months, while 17% (31/181) met the remission definition.
- For MADRS, 28% (57/202) of participants met the response definition at exit (LOCF analysis), and 20% (41/202) met the remission definition. For the observed population, 31% (57/181) met the response definition after 12 months, while 23% (41/181) met the remission definition.
- For IDS-SR30, 20% (40/201) of participants met the response definition at exit (LOCF analysis), and 13% (27/202) met the remission definition. For the observed population, 22% (39/180) met the response definition after 12 months, while 15% (27/180) met the remission definition.
- For the sustained response definition, (≥50% reduction in baseline HRSD24 symptoms at least once during the last quarter (months 9,10,11, or 12) and achieving at least a 40% reduction from baseline on at least two other of the HRSD24 assessments in the quarter), 27% (47/177) of participants met this criteria.

The publicly available FDA Clinical Memorandum comments that in the original protocol of June 2000, subjects’ HAM-D score were categorized by percent improvement: ≥75%; 50% to <75%; 25% to <50%; 0 to <25%; or, worsened (FDA Clinical Memorandum). In the 12 month completer population, 32% (56/177) improved ≥25%. In the revised statistical plan of September 2002, several additional measures were included, and are as follows (FDA Clinical Memorandum):

- The primary efficacy analysis was a repeated measures linear regression analysis performed on raw HAM-D scores during the first 12 months after initiation of stimulation on the 12 month completer population. The calculated endpoint was the average change in HAM-D per month over the first 12 months of stimulation, dividing the 12 months into quarters, the calculated as the average of the slopes across the four quarters, with each quarter having equal weight. The results showed these average changes in HAM-D per month: the 12 month completer population (N = 177), slope = 0.47 per month, $p < 0.001$; the evaluable population (N = 205), slope = 0.45 per month, $p < 0.001$; and the ITT population (N = 231), slope = 0.40 per month, $p < 0.001$.
- HAM-D categorical outcomes were examined as a ≥ 50% improvement in the score compared with baseline (response criteria) and the proportion of subjects with scores less than or equal to 9 (complete response criteria). For the 12 month completer population, 30% (52/174) met the response criteria, 17% (29/174) met complete response criteria.
- HAM-D categorical outcomes for the 12 month completer subjects were assessed over the last four visits of the year (months 9, 10, 11, 12) to evaluate which subjects had at least one visit with 50% or greater response and at least an additional two visits with at least a 40% or greater response. For the completer population, 27% (47/177) met this criterion.
- IDS-SR average change over 12 months also showed statistically significant improvement: for the 12 month completer population (N = 205, slope = 0.55 per month, $p < 0.001$); the evaluable population (N = 205, slope = 0.52 per month, $p < 0.001$); and the ITT population (N = 231, slope = 0.45 per month, $p < 0.001$).

- IDS-SR categorical outcomes were examined as a ≥ 50% improvement in the score compared with baseline (response criteria) and the proportion of subjects with a scores less than or equal to 14 (complete response criteria). For the 12 month completer population, 22% (38/173) met the response criteria, 15% (26/173) met complete response criteria.
- IDS-SR categorical outcomes for the 12 month completer subjects were assessed at months 9 and12 to evaluate which subjects had a 50% or greater response. For the 12 month completer population, 16% (27/174) met this criteria.

Adverse events were categorized similar to the acute phase study (FDA Clinical Memorandum). In this long-term phase, stimulation related adverse events captured only new adverse events related to stimulation not reported in the first three months (FDA Clinical Memorandum). These events included sudden unexplained death, hypotension, syncope (N=3), colitis (N=2), gastritis (N=2), weight gain (N=2), weight loss, arthralgia, joint disorder, myalgia, speech disorder, vocal cord paralysis, stridor, amblyopia, and deafness (N=2) (FDA Clinical Memorandum). Fifty one patients reported 96 serious adverse events including depression (N=62), convulsion (N=2), dizziness, drug dependence, manic depressive reaction, thinking abnormal, accidental injury, chest pain, overdose, peritonitis, sudden unexplained death, suicide attempt (N=7), and syncope (N=4) (FDA Clinical Memorandum). Rush et al stated, “During the 12-month study, 30 participants had worsening of depression sufficient to require hospitalization” and, “Two participants each made one suicide attempt (one coded by COSTART as an overdose) during the first 3 months of receiving VNS, and five participants made six suicide attempts over the ensuing 9 months of VNS (one participant made two attempts).” The FDA Clinical Memorandum notes a total of 34 total serious adverse events in 28 patients after the cut-off date of 10/10/2002. These events included, death, overdose, chest pain, suicide attempt (N=2), fibrillation atrial, syncope, depression (N=13), and pneumonia (FDA Clinical Memorandum). Three subjects had YMRS scores >15 without an adverse event reported (FDA Clinical Memorandum). For the entire study, six subjects had a manic/hypomanic reaction based either on clinical diagnosis (N=3) or YMRS scores (N=3) (FDA Clinical Memorandum). Suicidal ideation was assessed at 12 months of stimulation by an increase of at least two points in item 3 of the HAM-D: 3% (5/181) met this criterion, while 27% (48/181) decreased by at least two points (FDA Clinical Memorandum). Four deaths were reported during the D02 study: one before implantation (esophageal cancer), one death from suicide in the acute phase (treatment group), one death was listed as undetermined during the long-term phase, and the fourth death was after the long-term phase and was identified as nonspecific acute brain injury (FDA Clinical Memorandum).

In the FDA Clinical Memorandum, the sponsor included an additional analysis. They compared the results of D02 observational study to the results of a 2004 ECT study by Prudic et al. at seven community hospitals in the New York City area (Prudic et al. 2004). The comparison did not include all the patients studied in Prudic et al., but rather the subset that received ECT (N = 172/347). Inclusion/exclusion criteria of the two studies were not the same. The authors decided in this analysis to define response as ≥ 50% HAM-D improvement from baseline and remission was defined as a ≥ 60% HAM-D improvement from baseline to a score of 10 or less. In D02, 14% (29/205) evaluable patients were responders, 7% (14/205) met their complete response criteria at 3 months, and at 12 months, 27% (55/205) were responders and 15% (30/205) qualified as complete responders. For the ECT subset, at 3 months 58% (100/172) were responders with 44% (76/172) meeting the complete response criteria, and at 6 months, 41% (71/172) were responders and 20% (34/172) met complete response criteria (FDA Clinical Memorandum).

**Other Observational Studies**

**D01**

The D01 pilot study had as its objective to demonstrate the safety and efficacy of VNS for treatment of depression. This four-center case series study examined 60 patients followed over two years (Marangell et al. 2002; Nahas et al. 2005; Rush et al. 2000; Sackeim et al. 2001).

Inclusion and exclusion criteria are listed in Appendix C. To document a failed adequate treatment, the Antidepressant Treatment History Form was used, with information obtained from patient and family interviews, reports from treating physicians (level of documentation not specified), medical records, and pharmacy logs; however, the method of recruitment is not specified (advertisement, referral patients from primary care, patients in a tertiary care psychiatric clinic, etc). Nahas noted in a table footnote, “Other mood disorder treatments included mood stabilizers, psychostimulants, antipsychotics, anxiolytics, phototherapy, and other types of alternative treatments (e.g., St. John’s wort, flaxseed oil, and fish oil).”

The investigators chose to define an acute phase time frame (12 weeks after implantation) and long-term phase (>12 weeks after implantation up to 2 years). A total of 71 subjects enrolled, 11 discontinued prior to implantation, 60 were implanted, and 59 completed the acute phase (FDA Clinical Memorandum). Of the 11 who discontinued prior to implantation, 6 withdrew consent, and 5 failed to meet inclusion/exclusion criteria (Sackeim et al. 2001). One patient did not meet the criteria of scoring $\geq$18 on the HRSD-28 after implantation (HRSD-28 scores of 39 and 37 pre-implantation and scored 20 and 2 during the recovery period when the VNS device was not active) so was not included in the acute or long-term analyses but did enter the long-term phase study (Sackeim et al. 2001; FDA Clinical Memorandum).

Demographics revealed a mean age of 46.8 years (maximum age 63 ), 39/60 female, 59/60 Caucasian, mean duration of this episode 9.9 years, and age of onset of illness 28.7 years, with mean duration of illness 18 years (FDA Clinical Memorandum).

The acute phase included a 2 week recovery period after implantation, a 2 week stimulation adjustment period (adjusted to a “comfortable level”) and an 8 week period when the parameters of electricity delivery were held constant. In the long-term phase, acute phase responders and non-responders were followed differentially (FDA Clinical Memorandum). During the long-term phase, evaluations occurred at 6, 9, and 12 months, with device adjustment as needed, with responders having additional follow-up assessment at 4,5,7,8,10, and 11 months (FDA Clinical Memorandum). During the acute phase, no changes were made in medications or device parameters (after the 2 week adjustment period), though patients were allowed benzodiazepines (lorazepam up to 3mg/day). In the long term phase, device settings or concomitant medications could be changed based on investigator or primary physician judgment. Stimulation parameter settings were determined based on patient tolerance. Instruments used to report response during the acute phase: HAM-D, GAF, CGI, MADRS, BDI, YMRS and the SF-36 (FDA Clinical Memorandum). During the long-term phase, HAM-D, CGI, BDI, and IDS-SR occurred monthly for 12 months and then quarterly; MADRS, GAF, and YMRS occurred at months 6,9,12 and them quarterly (FDA Clinical Memorandum). The SF-36 occurred at 12 months and then again a year later (FDA Clinical Memorandum).

Of 60 subjects entering the long term phase, 8 subjects discontinued VNS and 7 had their devices removed (leaving 52 subjects as of 10/29/2002). Included in the 8 subjects, were 2 subjects who died, and one subject withdrew consent (FDA Clinical Memorandum). Nahas reported this as, “At the 24-month follow-up, 53 participants remained implanted with the VNS device. Ratings were available at the 24-month follow-up for 42 of the 53 implanted patients.” Then Nahas states, “Of the original 59 participants, 6 were no longer implanted at the 24-month follow-up: 2 had died, and 4 had the device explanted owing to lack of efficacy,” and, “Six patients had been set to zero mA between 15 and 24 months (including 1 patient who was explanted at 24 months).” “Patients varied considerably in the type and intensity of concurrent pharmacological treatment they received during the acute VNS trial” (Sackeim et al. 2001). While in the acute study, patients were taking a median of 1 (range of 0 to 4) antidepressant medications and a median of 4 (range 0 to 10) mood disorder treatments (Sackeim et al. 2001). Some subjects did have changes during the 4 weeks prior to the first visits (6) and minor changes during the acute phase (8) (FDA Clinical Memorandum). Four subjects started a new antidepressant medication during the period from visit 5 to visit 12, due to worsening depression (FDA Clinical Memorandum). Three subjects with bipolar depressive disorder and two subjects with major depressive disorder received lithium during the acute phase (FDA Clinical Memorandum). While no subjects received ECT during the acute phase, 3 received ECT during the long-term study (FDA Clinical Memorandum). Adverse events were recorded differently for the acute and long-term phase (FDA Clinical Memorandum).

During the acute phase, all adverse events were recorded (FDA Clinical Memorandum). “During the long-term phase, only adverse events that were considered by the investigator to be possibly, probably, or definitely related to either implantation or stimulation were reported” (FDA Clinical Memorandum). All implanted subjects reported at least one treatment related adverse event (FDA Clinical Memorandum). The most common events included device site pain, headache, incision pain, neck pain, pain, dysphagia, increased cough, dyspnea, and voice alteration (FDA Clinical Memorandum). The FDA Clinical Memorandum noted 77 serious adverse events reported after implantation and included worsening depression (N=34), suicide attempt or overdose (N=12), mania (N=2), and agitation (N=2). Nahas reported, “…40 serious AEs (SAEs) involving 25 participants occurred,…” and, “These 40 AEs included 3 for suicide attempts, 10 for worsened depression, 1 for dysphoria, 2 for a manic episode, 1 for agitation, and 1 for CNS toxicity. All other SAEs were not psychiatrically related. No SAE was thought to be device related.” Other events included syncope, deep thrombophlebitis, wound infection, neuroleptic malignant syndrome, pain, and convulsion (FDA Clinical Memorandum). One death was in the long-term phase (multiple organ failure) (FDA Clinical Memorandum). Another subject died from lung cancer after withdrawal from the study. The primary efficacy endpoint was response defined as a $\geq$ 50% decrease in the HAM-D score at post-treatment acute phase exit as compared with the baseline period, and then evaluated again at 1 and 2 years. Relapse and recurrence were not defined.

At the 10 week follow-up, 31% (18/59) of study subjects had a 50% reduction in the primary outcome of the HRSD-28 (FDA Clinical Memorandum; Sackeim et al. 2001; Nahas et al. 2005). At one year, 45% (25/55) met response criteria, at two years 43% (18/42) (FDA Clinical Memorandum). Complete responders were defined by the Rush and Sackeim as HAM-D score $\leq$ 10 (which was defined as remission by Marangell and Nahas in the long-term phase) and included 15% (9/59) at acute phase exit, 27% (15/55) at 1 year and 21% (9/42) at 2 years (FDA Clinical Memorandum). Alternately, Nahas reported the results based on LOCF analyses, with the HAM-D 28 response rate of 31% (18/59) at 3month, 44% (26/59) at 1 year, and 42% (25/59) at 2 years, with remission (HAM-D 28 $\leq$ 10) 15% (9/59) at 3 months, 27% (16/59) at 1 year, and 22% (13/59) at 2 years. Nahas also reported as a post-hoc analysis the durability of response which the authors define as, “calculating the percentage of patients who met modified response criteria at the 12- and 24-month time points, and who had met a priori response criteria at earlier time points (3-month or 12-month). In these post hoc analyses, individuals who were responders at the earlier time point and who had at least 40% improvement in HAM-D-28 scores relative to baseline at the subsequent follow-up were classified as showing sustain response.” Nahas noted, “Although response rates were not significantly different at 12 and 24 months, individual responses varied considerably over time …” Outcomes from secondary efficacy variables included the MADRS, CGI, GAF, BDI, IDS-SR, YMRS, subject diary, and SF-36 (FDA Clinical Memorandum). Sackeim noted response rates at the four sites: four of eleven patients (36%); one of twelve patients (8%); four of thirteen patients (31%); and nine of 23 patients (39%).

Marangell, Rush, and Sackeim concluded that “Response rates were highest among patients who showed fewer unsuccessful adequate antidepressant treatment trials” (Marangell et al. 2002). Based on the 24 month LOCF analysis in Nahas, the authors did not find an association between greater treatment resistance (ATHF score) and response.

**D04**

This sponsor funded observational study did not involve the VNS treatment. The data was first published (George et al. 2005) as a comparison (called treatment as usual, TAU) for the D02 study, but, “The TAU group had not originally been intended to serve as the primary benchmark for the VNS + TAU group; it was intended to describe health care costs.” The published description of the study (Dunner et al. 2006) stated, “The study was designed to assess (1) the clinical characteristics of a population with TRD; (2) the percentage of patients who met response and remission criteria at each measurement occasion; (3) the changes in functional health and well-being that occur over time.”

This prospective study followed patients with treatment resistant depression receiving standard of care. Standard of care was defined as the treatment plan the physician and patient chose to follow, including: medications (antidepressants, stimulants, thyroid hormone, lithium, atypical antipsychotics, and anticonvulsants), psychotherapy, bright light therapy, or ECT. The authors noted there are many obstacles to treatment of MDD, including: misdiagnosis, undertreatment, lack of treatment, and patients who discontinue treatment. They also noted, “In addition, other treatment obstacles include the lack of a standardized definition of TRD, a poor understanding of the clinical characteristics of patients with TRD, and limited evidence for how to best treat this population.” The study was to enroll 130 patients to ensure data on 80 subjects at 12 months, with a goal of 100 subjects with 12-month of data across 15 sites (FDA Clinical Memorandum). Referral sources were from community psychiatrists or were under the care of the investigator at the study site. Patients were cared for by their referral provider (with study assessment every 3 months) or the investigator. Data were collected at 13 sites (12 sites overlapped with D02 sites). Clinical, quality of life, and economic outcomes were assessed at baseline, 3, 6, 9, 12, 15, 18, 21, and 24 months (FDA Clinical Memorandum). During the first 12 months the MADRS, CGI, IDS-SR, YMRS, and SF-36 assessments were reviewed quarterly. HAMD 24 assessments were performed at baseline and 12 months, then quarterly (FDA Clinical Memorandum). Patient safety information was not systematically collected.

Inclusion and exclusion criteria based on Dunner et al. 2006 and the FDA Clinical Memorandum are listed in Appendix C.

**Table 3:** D04 Subject Tracking (FDA Clinical Memorandum)

| Tracking Point | Subject Number |
|---|---|
| Enrollment | 138 |
| Discontinued | 11 |
| Baseline data only | 3 |
| | |

| Tracking Point | Subject Number |
|---|---|
| Evaluable Subjects | 124 |
| Not 12 month completers | 12 |
| 12 month completers | 112 |

Eleven patients did not meet study inclusion criteria at baseline (no details available). Two patients withdrew at baseline and one had no post baseline assessment. At 12 months, 112 patients were evaluable, and 103 were evaluable at 24 months. Of the 21 patients who did not complete the study, 13 withdrew consent (no details), 3 were excluded because of significant noncompliance, and 5 were withdrawn wither due to investigator decision (no details were given) or loss to follow-up. Mean age was 45.5 (no range given), 68% female, 90% Caucasian, 12% bipolar I or II, 75% recurrent depression.

Data analysis included a repeated measures linear regression performed on the raw IDS-SR scores during the first 12 months on evaluable subjects (N =124) (FDA Clinical Memorandum). Subjects were assessed for a $\geq$50% improvement in IDS-SR at the last two measured quarters (the primary outcome). Remission was defined by the investigators as an IDS-SR-30 score of $\leq$ 14. Response and remission were determined at each visit and were defined by the symptom severity for the prior 7 days. Generalized estimating equations approach was used to analyze change over time, with baseline values as a covariate. The FDA Clinical Memorandum noted, “Of the 12 month completer population, 4% (5/109) reportedly met the proposed criteria.” Dunner et al stated, “The 12 and 24 month IDS-SR30 response rates were 11.6% (13/112) and 18.4% (19/103), respectively.” Participants who met the 12 and 24 month study definition of remission were 4/112 and 8/103, respectively (Dunner et al. 2006). For CGI, 12/101 of evaluable patients were rated as much improved or very much improved at 12 months (FDA Clinical Memorandum). The authors stated, “Changes in quality-of-life measures were minimal, with SF-36 subscale scores remaining predominately below average for the duration of the study.” Dunner et al do not report MADRS, CGI, YMRS, or HAMD 24 outcomes.

**D05**
D05 is a videotape assessment of the D02 study to examine inter-rater reliability for the depression assessments.

**D06**
D06 is a pilot study of VNS in patients with rapid cycling bipolar disorder, for which published data is not available.

**Other Studies**

**Comparison of D02 participants to D04 participants**

The objective of this analysis (George et al. 2005) was to compare the D02 (participants receiving VNS therapy in addition to other therapies for depression) outcomes to D04 (participants receiving therapies for depression) outcomes. The inclusion and exclusion criteria for D02 study and D04 study were not the same (Appendix C) (FDA Clinical Memorandum).

For the first 3 months of the D02 study, subjects were required to maintain a stable mood disorders medication regimen and ECT was not allowed. After 3 months of the D02 study, changes to mood disorders medication and additional treatments such as ECT were permissible. In D04, the definition of standard of care was whatever treatment strategy the physician and subject chose to follow. Neither study specified any criteria for the added or increased use of non-VNS antidepressant treatments (FDA Clinical Memorandum).

IDS-SR was the primary endpoint for this comparison, using a repeated measure linear regression analysis of the raw IDS-SR scores. George et al stated, "The VNS + TAU participants in this report are the evaluable 12-month sample (n = 205) described by Rush et al (2005b)." The D04 population had 124 participants. There was no reference to enrolled, discontinued, how many subjects provided enough data to be evaluated, and how many were completers (this analysis was published before the D04 trial was published). Some of the D02 and D04 sites overlap (D04 included 13 sites, 12 also participated in D02), and the majority of D04 participants enrolled after D02 was closed (FDA Clinical Memorandum). The FDA Clinical Memorandum noted, "Although both Study D-02 and D-04 were available to enroll subjects at similar time periods, almost all D-04 subjects enrolled into the study after D-02 was closed for enrollment." A baseline demographic comparison was done between 205 D02 participants and 124 D04 participants. These reported parameters were statistically similar ($p \geq 0.05$): age; gender; diagnosis (unipolar or bipolar); unipolar type; length of current MDE; participants having chronic ($\geq$ 2 y) current MDE; number of failed adequate trials in current MDE by ATHF; number of failed adequate trials in current MDE per year of MDE; age at onset of first symptoms of depression, mania, hypomania; age of definitive diagnosis of any mood disorders; duration of illness; length of time since definitive diagnosis; length of time between onset and definitive diagnosis; number of suicide attempts in lifetime; number of suicide attempts within past 12 months; treatment-induced hypomania or mania; number of prior hospital admissions for mood disorders in lifetime. These reported parameters were statistically different ($p < 0.05$): ethnic origin (Caucasian: D02, 97%; D04 90%); received ECT, lifetime (D02: 53%, D04 26%); Received ECT, current MDE ( D02 35%; D04 12%); number of lifetime episodes of depression (0-2: D02 24%, D04 25%; 3-5: D02 34%, D04 29%; 6-10: D02 27%, D04 15%; >10: D02 9%, D04 26%; unknown: D02 5%, D04 6%). A propensity score was derived from the previously listed parameters, and included baseline HRSD 24 and IDS-SR30 score and was grouped into a five-level category.

The FDA Clinical Memorandum stated, "a statistically significant difference ($p < 0.001$) was observed in the estimated IDS-SR raw scores per month between D02 and D04 at 12 months (-0.397 estimated average difference per month)". George et al presented the results as a repeated-measures linear regression model, in which, "The model-estimated differences (SE) in the IDS-SR30 total score at the end of 3,6,9 and 12 months were -1.19 (.29), -2.38 (.58), -3.57(.87), and -4.76 (1.16) points, respectively." A note was made in the figure representation of this model that "All VNS+TAU measures within a quarter were assigned to the end of the quarter. This model adjusted for baseline IDS scores, propensity score, and site." To examine the impact on concurrent antidepressant treatments upon the long-term outcomes, an asymmetric analysis using the primary repeated measures linear regression analysis of IDS-SR scores and censoring the D-02 participants rating scores for the concurrent antidepressant treatment changes (LOCF approach) was presented (FDA Clinical Memorandum). The FDA Clinical Memorandum reported:

"The results (not specified in the Sponsor's original clinical protocol or the Sponsor's revised clinical protocol) reported the following: D02, D04 Primary Analysis Comparisons, *After* Censoring Scores for Concomitant Antidepressant Treatment Changes-If a subject added or increased a concomitant antidepressant treatment (D02 only; D04 standard of care was defined as whatever treatment strategy the physician and subject chose to follow), and their subsequent IDS-SR scores were not used (i.e., a censored analysis employing a last-observation-carried-forward approach) in a revised repeated measures linear regression analysis (Table 2), the difference observed in the estimated IDS-SR raw scores per month between the D02 and D04 evaluable populations at 12 months (-0.183), i.e., the average amount of improvement in the IDS-SR score per month in D02 and D04, was not statistically significantly different from standard of care ($p>0.05$)."

George et al stated, “With an ANCOVA model adjusting for the same covariates, the estimated difference (SE) in scores between the groups was -6.2(1.7) points at 12 months, and -5.1(1.6) points, last observation carried forward (LOCF).” The authors also concluded, “Thus the observed baseline demographic and illness characteristics between the two groups likely did not contribute significantly to the difference in IDS-SR30 outcomes between the VNS + TAU and TAU groups.”

A number of secondary endpoints were compared. For IDS-SR rating score meeting the study definition of response (percentage of subjects reporting $\geq$ 50% decrease in the raw IDS-SR score between baseline and 12 months) and complete response (percentage of participants with an IDS-SR raw score of less than 14 at 12 months)(FDA Clinical Memorandum):

- For 12 month evaluable populations, 22% (39/180) of the D02 and 12% (13/112) of the D04 met the definition of response (0 = 0.029). After censoring scores for concurrent antidepressant treatment changes for the 12 month evaluable populations, 18% (32/180) or the D02 and 3% (13/112) of the D04 met the definition of response (p = 0.085). The authors stated this as LOCF response rates, D02 19.6% (n=204); D04 response 12.1% (N=124); p=0.002 (George, Rush 2005)
- For the 12 month evaluable populations, 15% (27/180) of the D02 and 4% (4/112) of the D04 met the definition of complete response (p = 0.006). After censoring scores for concurrent antidepressant treatment changes for the 12 month evaluable populations, 7% (12/180) or the D02 and 4% (4/112) of the D04 met the definition of complete response (p = 0.048). The authors stated this as LOCF remission rates, D02 13.2%; D04 3.2% (n=124); p = 0.007.

For the HAM-D rating score meeting the study definition of response (percentage of subjects reporting $\geq$ 50% decrease in the raw HAM-D score between baseline and 12 months) and complete response (percentage of participants with a HAM-D raw score of 9 or less at 23 months)(FDA Clinical Memorandum):

- For 12 month evaluable populations, 30% (54/181) of the D02 and 13% (13/104) of the D04 met the definition of response (p = 0.003). After censoring scores for concurrent antidepressant treatment changes for the 12 moth evaluable populations, 20% (36/181) or the D02 and 13% (13/104) of the D04 met the definition of response (p = 0.118). The authors reported this as LOCF response rates, 26.8% (n=205) for D02; 12.5% (n=104) for D04; p = 0.011.
- For the 12 month evaluable populations, 17% (31/181) of the D02 and 7% (7/104) of the D04 met the definition of complete response (p = 0.031). After censoring scores for concurrent antidepressant treatment changes for the 12 month evaluable populations, 8% (15/181) or the D02 and 7% (7/104) of the D04 met the definition of complete response (p = 0.757). The authors reported this as LOCF complete response rates (%), 15.6 (n=205) for D02; 6.7% (n=104) for D04; p = 0.059.

The unipolar and bipolar patients were analyzed as subgroups (FDA Clinical Memorandum).

**Table 4:** Unipolar patients only (FDA Clinical Memorandum)

| | | | |
|---|---|---|---|
| Study Outcome | D02 | D04 | p-value |
| **IDS-SR score** | N=163 | N=97 | |
| Average change per month | -9.5 | -4.7 | 0.001 |
| Average change per month LOCF calculation | -8.8 (N=184) | -5.1 (N=109) | 0.011 |
| Number of responders | 34 | 12 | 0.084 |
| Number of complete responders | 25 | 4 | 0.014 |
| | N=181 | N=104 | |

| | | | |
|---|---|---|---|
| **HAM-D score** | | | |
| Average change per month | -8.1 | -4.8 | 0.013 |
| Average change per month LOCF calculation | -7.1 (N=185) | -4.8 (N=91) | 0.070 |
| Number of responders | 49 | 11 | 0.005 |
| Number of complete responders | 27 | 7 | 0.096 |

**Table 5:** Bipolar patients only (FDA Clinical Memorandum)

| | | | |
|---|---|---|---|
| Study Outcome | D02 | D04 | p-value |
| | N=17 | N=15 | |

| | | | |
|---|---|---|---|
| **IDS-SR score** | | | |
| Average change per month | -12.6 | -3.7 | 0.703 |
| Average change per month LOCF calculation | -13.8 (N=20) | -3.7 (N=15) | 0.976 |
| Number of responders | 5 | 1 | * |
| Number of complete responders | 2 | 0 | * |
| **HAM-D score** | N=17 | N=13 | |
| Average change per month | -9.5 | -5.6 | |
| Average change per month LOCF calculation | -9.7 (N=20) | -5.6 (N=13) | 0.340 |

| | | | |
|---|---|---|---|
| | | | |
| Number of responders | 5 | 2 | * |
| Number of complete responders | 4 | 0 | * |

* cell size too small to calculate a p-value

Participants' IDS-SR and HAM-D scores were also examined in other ways. The IDS-SR scores were compared for a 50% improvement or better at the last two measured quarters of the first year of VNS therapy (visits 9 and 12 month). For 12 month completer populations, 15% (27/177) of D02 and 4% (5/112) of D04 met this criterion (FDA Clinical Memorandum). For the evaluable populations, 13% (27/205) of D02 and 4% (5/124) of D04 met this criterion (FDA Clinical Memorandum). Participants HAM-D raw scores were compared for improvement from baseline to 12 months. For 12 month completer participants, the D02 (N=180) average 12 month score was an 8.2 point decrease from baseline while D04 (N=104) had a 4.9 decrease (statitistically significant). For the LOCF analysis the D02 (N=205) average 12 month score was a 7.4 point decrease from baseline while D04 (N=124) had a 5.0 decrease. Additional analyses were performed with other collected scores from other scales. For CGI, 36% (66/181) participants were rated as much improved, and 12% (12/101) were rated as much improved for D04 (LOCF analysis: D02 68/200, D04 12/101) (FDA Clinical Memorandum). The authors reported this as 36.5% (n=181) for D02; 11.9%(N = 101) for D04; $p < 0.001$ (LOCF analysis D02 34.0% (N=200), D04 11.9% (N=101), $p < 0.001$) (George et al. 2005). The FDA Clinical Memorandum stated, "the MOS-SF36 had numerically greater changes in vitality, social functioning, role functioning-emotional, and mental health. No statistical comparisons were performed between D02 and D04." Adjustment for multiple comparisons was not done.

The authors attempted to investigate confounders. To determine if site differences contributed to the differences between D02 and D04, the authors reexamined the outcomes using data from the 12 overlapping sites only. The primary results were stated as, “Analysis with the primary repeated-measures linear regression model remained significant and resulted in a linear study effect of a difference between groups of .32 IDS-SR30 point per month [SE=.10, t(862) = 3.16, p=.002].” To determine if mood medication changes brought about response, they examined the addition or increase in antidepressants and mood stabilizers in responders and nonresponders from both D02 and D04. D02 responders had fewer dose increases or medication additions (56%) as compared to nonresponders (77%), while for D04 more responders had increases or additions (92%) than did nonresponders (80%). The authors created a plot and referred to it as a sensitivity analysis: “Plot of 30-item inventory of depressive symptomatology-Self-Report (IDS-SR) mean scores for participants receiving vagus nerve stimulation plus treatment as usual, with and without medication changes (n= 205).” The conclusion by the authors was, “Symptom reductions in VNS+TAU were largely attributable to participants without a medication change or increase during the 12 months.”

**4. MCAC**

No MCAC was held for this topic.

**5. Evidence-based guidelines**

The 1993 AHRQ guidelines for Detection and Diagnosis of Depression in Primary Care and Treatment of Major Depression are listed as outdated for current medical practice on the internet.

APA treatment guidelines for patients with major depressive disorder (2000) are, “in some instances based on data distilled from randomized prospective clinical trials, while in other areas they are based on individual case reports along with the collective experience and judgment of well-regarded senior psychiatrists.” In this guideline, treatment resistant depression is not listed. “Failure to respond to pharmacotherapy in the acute phase,” is the title of a section. The guideline states, “ For patients whose treatment failure is not readily attributable to inappropriate diagnoses, poor adherence, or complicating conditions, a variety of therapeutic options are available, including maximizing the initial treatment, switching to another non-MAAOI, agent, augmenting antidepressant medications with other medications or psychotherapy, using an MAOI, and ECT. Empirical data concerning the relative efficacies of these strategies are limited.” VNS is not mentioned. In Fochtmann and Gelenberg’s Guideline Watch for the Practice Guideline for the Treatment of Patients with Major Depressive Disorder, published September 2005, this statement is included in the topic of somatic treatments: “Although other somatic treatments, including repetitive transcranial magnetic stimulation, magnetic seizure therapy, and vagal nerve stimulation, have also been studied over the past 5 years, evidence is not yet sufficient to recommend their use in routine clinical practice.”

**6. Professional Society Position Statements**

None

**7. Expert Opinion**

Dr. John A. Rush, Dr. Harold Sackeim, Dr. Michael Thase, Dr. Mark George, Dr. Dunner, Dr Mustafa Hussain, were identified as vagus nerve stimulation experts for the treatment of resistant depression. These commenters provided evidence, much of which was papers they coauthored. All supported the use of VNS in TRD.

**8. Public Comments**

**Initial 30-day comment period**

During the initial 30-day comment period, CMS received 1,843 comments. Of the comments received, 1,274 were submitted through the public comment process via CMS coverage website and another 569 were submitted via US mail or to the analyst's e-mail.

Of the total 1,843 comments, 1,831 commenters were in favor of amending the current VNS policy to include treatment-resistant depression as a covered indication. Seven commenters were against changing CMS' current policy. Five commenters expressed no indication for coverage.

**A. Comments without evidence**

Patients

CMS received 618 comments from the patient population. All patients who commented during this comment period are in favor of expanding the coverage policy to include TRD as a covered indication. 122 patients who had undergone VNS surgery for TRD submitted comments. Also, nine patients who had VNS implanted for epilepsy commented. Of the patient commenters, more than 500 reported being diagnosed with depression, seven patients reported being bi-polar and/or depressed, more than twenty reported they were suffering from anxiety and/or depression, and nine patients did not report a diagnosis. It was unclear how many patients suffer from TRD. However, many commenters mentioned they suffered from comorbidities such as panic disorder, obsessive compulsive disorder, psychosis, post traumatic stress disorder, anorexia, bulimia, generalized anxiety disorder, and other medical conditions, such as diabetes. Commenters also mentioned they were experiencing social difficulties, such as unemployment, marital difficulties (including divorce), difficulties in caring for children (particularly single parents), and history of sexual abuse. Some expressed suicidal ideation and violent thoughts (including homicide).

Many patients included statements such as, "I am sending the enclosed information regarding my lifetime struggles with Treatment Resistant Depression." One patient commented that he received a PhD but has not been able to pursue any postdoctoral work because of depression. Another commenter said, "In the last ten years, I have twice tried working as a cashier, and even that was too much for me." Other typical comments included ones such as, "My inability to work stems not only from my illness but also from the medications and electroconvulsive therapy which I endure to keep my depression under control." Many patients stated that they have failed from 20 to 30 antidepressant treatments.

Patients expressed confidence in the FDA decision with statements such as, "The FDA has determined that VNS Therapy has a level of safety and effectiveness demonstrated in data showing 'significant' and 'sustained' improvement over time." For patients who have already received VNS therapy, some claim to have had almost immediate total response, while others claim to be improving very slowly. For patients who have received VNS therapy, statements such as this were noted, "I feel 100% better and have been able to stop taking so much medication."

Based on the comments received, the number of patients receiving Medicare through disability or due to age is unclear. Many commenters stated they were appealing a denial of coverage determined by a private insurer, the number of patients receiving medical benefits through insurance other than Medicare was also unclear. In addition many of these same commenters state that they are hoping Medicare will approve this service so other insurance carriers will follow Medicare's lead.

It is unclear how many patients and physicians were solicited by Cyberonics resulting from the company's attempt to encourage public comments in this issue. Many commenters stated that they were submitting a comment in response to a solicitation made by Cyberonics.

Physicians

The table below indicates the number of physicians who commented (without evidence) during the initial comment period. (This table includes each physician's reported level of VNS experience). Many physicians enclosed their Curriculum Vitae (CV), as suggested by Cyberonics.

| Physician Type | Total Number | Experience with VNS for TRD | Experience with VNS for Epilepsy | Experience with both TRD/Epilepsy | No reported experience |
|---|---|---|---|---|---|
| Psychiatrists | 252 | 162 | 8 | 0 | 82 |
| Surgeons | 50 | 26 | 15 | 5 | 4 |
| Neurologists | 45 | 4 | 36 | 3 | 2 |

| Physician Type | Total Number | Experience with VNS for TRD | Experience with VNS for Epilepsy | Experience with both TRD/Epilepsy | No reported experience |
|---|---|---|---|---|---|
| | | | | | |
| Other | 49 | 22 | 7 | 0 | 20 |

All physicians (except one) who submitted comments were in favor of extending coverage. Some physicians described their experience with VNS with comments such as, “So far we’ve had nearly 100% response and the response has been dramatic in most of our patients and these were all people who had failed multiple courses of depression medicines, multiple electroconvulsive therapies and multiple hospitalizations for depression.” Many neurosurgeons said, “While I generally only see these patients for a short period of time during their preoperative and postoperative period, many patients report to me that their mood has significantly improved.” Others commented on FDA approval with statements such as, “The FDA was sufficiently convinced of the successfulness of this therapy in this population to give their stamp of approval for it.” Others expressed that as the FDA had thoroughly reviewed the device, the device should be covered by insurance. Some agreed with the FDA’s indications, others did not, and included comments such as, “In my opinion requiring ECT as a prerequisite for all VNS treatment of depression would be ‘coercion’.”

One physician was against extending coverage. He commented, “Given the many, many less expensive interventions that are not supported by Medicare and most other payers, payment for epilepsy or depression is an insult.” He continued to say, “The manufacturer has expended substantial money lobbying and a very lucrative payment schedule has been created for providers who implant and adjust devices.”

Other Health care professionals/organizations

Of the comments without evidence, CMS received 109 comments from other healthcare professionals (103 favored coverage and six opposed coverage). This group included nurses, advocacy groups, social workers, health plans, and others.

CMS received a comment from the Epilepsy Foundation. They stated, “As you know, VNS Therapy recently received approval for the FDA and has demonstrated substantial and compelling clinical evidence of long-term effectiveness in patients with chronic and recurrent TRD.” They further added, “The Epilepsy Foundation therefore is of the opinion that because mood disorders are such a serious, concomitant condition of epilepsy, CMS should provide national coverage for all FDA approved treatments for mood disorders.” The President of the Depression and Bipolar Support Alliance commented, “I hope the medical community will do more to help understand and effectively treat this illness. VNS offers hope - the possibility of recovering a full and productive life where we can contribute to the world.”

Commenters who were against extending coverage included: Geisinger Health Plan, American Managed Behavioral Healthcare Association (AMBHA) and its member companies, and America’s Health Insurance Plans (AHIP). One of these commenters stated that “a decision to cover this service would be premature and unfounded.” This same commenter argued that, “VNS for treatment of TRD is of unproven benefit and safety.” Another commenter argued, “I believe the proposal to approve the VNS as an effective treatment for depression flies in the face of the available evidence, and that it is not appropriate or justified.” The AHIP commenter stated that the FDA lowered its threshold for evidence and effectiveness in the case of VNS for TRD. She also added, “This concept of allowing approval under the auspices of gathering further data is central to CED, but there is a serious danger of misusing the CED mechanism in a case such as VNS when the evidence is in such early stages and cannot provide even minimal assurance of effectiveness, longer-term safety, and avoidance of potential risk.”

Professional Societies

The American Psychiatric Association (APA) submitted a comment in support of including TRD as a covered indication for VNS.

The President of the APA commented that, “TRD is a serious medical condition and while there is no official definition of TRD, the term is used in clinical psychiatry to describe cases of major depressive disorder that do not respond to typical modes of treatment (i.e.: psychotherapy and common antidepressants such as SSRIs).” He concluded, “Given the existence of this population of Medicare patients with TRD, who have a clinical history that clearly shows that there are no other medically beneficial treatment alternatives available to them, and given the FDA’s approval of the safety and effectiveness of VNS for treatment of major depressive disorder, it is our medical opinion that VNS is, as defined by §1862 (a)(1)(A) of the Social Security Act, ‘reasonable and necessary’ for the treatment of those Medicare patients with treatment-resistant depression.”

General public

CMS received 685 comments from the general public. All comments submitted by the general public favored coverage. These comments included input from family members, friends, and others from the general public. Letters included comments such as, “I am the son-in-law of a woman who is in great need of the hope, relief, and aid that the Vagus Nerve Stimulation Therapy offers.”

Many commenters also expressed concern relating to parity of coverage. For instance, many said that since CMS covers VNS for epilepsy patients, coverage should be extended to include the mental health community. Commenters also said that coverage of mental health treatments is not treated as fairly as other health issues such as cancer treatment. For example, one commenter stated, “Somebody with cancer or diabetes has so many treatment options that are recognized and approved by many insurance companies as well as Medicare and Medicaid Services (CMS).”

One patient had over 170 friends and family members comment on her behalf. Many commenters mentioned that VNS implantation is more cost effective than current TRD-related treatment modalities.

**B. Comments with evidence**

CMS received a total of 70 comments that referred to evidence to support coverage of VNS for TRD. We also received several comments that were not in favor of Medicare coverage. Articles and information provided as evidence included: studies and information already mentioned; studies of depression and treatments for depression other than VNS, including the NIH funded STAR-D trials; physiologic studies; VNS in rats; review articles; studies on the cost of depression; a small case series study (11 patients) funded by the sponsor; a case report of VNS; the use of VNS in epilepsy; other non-peer reviewed information such as posters/ abstracts, newspaper articles, and letters to the FDA; and other information unrelated to depression or VNS (hematopoietic cell transplantation). Physiologic studies and animal studies generally do not provide evidence of clinical benefit that is particularly pertinent to CMS. Review articles do not provide additional information. Non-peer reviewed information generally is accorded less weight than peer reviewed material. Studies of treatments for depression other than VNS provide no relevant evidence concerning the benefit of VNS. These references are included in the final reference list.

CMS received two comments with evidence that opposed extending coverage. Public Citizen's Health Research Group commented, "Indeed, while under the Food, Drug, and Cosmetic Act, a device must be proved safe and effective to gain FDA approval, the Social Security Act provides for reimbursement under Medicare only if the device is 'reasonable and necessary.' In our view, neither standard has been met, and for this reason today we are also filing a petition with the FDA urging the reversal of FDA's scientifically meritless previous decision to approve VNS." They further add, "Even the full-court press of misleading advertising, training sessions in its use for physicians presentations at the American Psychiatric Association annual meeting, case managers to help secure reimbursement for individual patients, abuse of FDA employees, misleading clinical trial write-ups, ghost-written review articles and company-generated favorable local media coverage cannot disguise what is lacking and what insurers are increasingly realizing: There are no convincing data of the device's effectiveness, let alone, in CMS terms, that it is 'reasonable and necessary'."

In addition, the Senate Finance Committee provided us its report regarding FDA's approval of VNS for TRD entitled, "Committee Staff Report to the Chairman and Ranking Member, Review of FDA's Approval Process for the Vagus Nerve Stimulation Therapy System for Treatment-Resistant Depression." The Senate Finance Committee requested that CMS take this report into consideration as we make a determination on whether VNS should be covered.

The Blue Cross Blue Shield Association commented with the inclusion of the results from its Technology Evaluation Center (TEC) assessment. This commenter stated that the available evidence for VNS for TRD was reviewed in 2005 and the Association determined that VNS did not meet the TEC criteria regarding the effect of VNS on health outcomes. In 2006, the TEC re-assessed VNS for this indication but its conclusion remained the same. The TEC stated, "The clinical trials reviewed report weak evidence that does not demonstrate efficacy. This evidence does not permit conclusions regarding the effect of VNS therapy on health outcomes or its effect compared with alternative therapies. For this reason, VNS for treatment-resistant depression should not be considered a reasonable and necessary service under the Medicare program."

**VIII. CMS Analysis**

National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally under title XVIII of the Social Security Act § 1869(f)(1)(B). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions, the expenses incurred for items or services must be "reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member." § 1862(a)(1)(A).

CMS focused on this general question:

Is the evidence sufficient to conclude that, in the Medicare population, vagus nerve stimulation will improve health benefits for individuals with treatment resistant depression?

Depression is a common disorder that is a major cause of morbidity worldwide. Some have suggested that depression is not a single disease, but “a heterogeneous syndrome comprised of numerous diseases of distinct causes and pathophysiologies” (Nestler et al. 2002). The diagnosis of this condition is based on subjective symptoms that can be variable. While there appear to be many effective treatments, there is little to guide clinicians to treatment selection in individual patients. There is little to guide clinicians after initial treatment failure. Methods used to measure treatment progress lack standardization, and are inadequate to reflect the complexity of the disorder. Though depression has been recognized since ancient times, patient care basics such as diagnosis and treatment remain very challenging.

About 10 to 30% of patients with depression fail to respond to treatment (Cadieux 1998). There are varying definitions of treatment resistance, response, and remission, in the published literature; therefore, determining the number of people who are actually resistant to treatment is difficult (Keller 2005). Nonresponse may come from a variety of causes or contributing conditions: pharmacokinetic or pharmacogenomic factors for which there is wide interindividual variation (Fleck & Horwath 2005); treatment adherence (“Depressed patients, who typically feel hopeless and lack motivation, often discontinue treatment. In one study, even when patient adherence was monitored through monthly telephone interviews, 53% of patients discontinued treatment within 6 months”) (Zajecka 2003); comorbid medical conditions that cause depression; comorbid psychiatric conditions; substance abuse; psychosocial conditions; and, drugs that cause or worsen depression. Lastly, inadequate treatment is common, with many patients being left undertreated and with residual symptoms. A strong need exists for quality improvement in the area of treatment for depression (Zajecka 2003);(Kessler et al. 2003).

The vagus nerve has influence over widespread brain areas (Groves & Brown 2005). Vagus nerve stimulation is used in treatment of epilepsy and is under investigation for use in a variety of other conditions including obesity and cognitive disorders. Improvement in mood was observed in some patients being treated for epilepsy with VNS therapy. In studies of MDD, it was hypothesized for use in patients with chronic or recurrent depression defined as patients who had failed two antidepressant treatments (treatment resistant depression). It was approved by the FDA for patients who had failed four antidepressant treatments. Additional requirements have been suggested, by the requestor, for coverage in those previously treated with or refused treatment with ECT or who have been previously hospitalized for depression. These various indications without scientific basis illustrate that the term treatment resistant depression lacks a standard definition that has been scientifically validated, and appears to be subject to arbitrary interpretation. The National Clinical Practice Guideline for Depression in Primary and Secondary Care (2004) states an important point: “The term ‘refractory depression’, used to describe depression that has failed to respond to two or more antidepressants at an adequate dose for an adequate duration given sequentially, is not especially helpful. It does not take into account depressive subtypes, makes no distinction between treatment resistance, chronicity, relapse or recurrence, and fails to take into account what psychosocial factors may be preventing recovery or indeed whether the patient has had an adequate course of an appropriate psychotherapeutic treatment (Andrews & Jenkins, 1999)”.

The effectiveness of treatments for depression cannot be conclusively judged from case series data, due to regression to the mean (the waxing and waning of symptoms), spontaneous remission (which is known to occur), and placebo response, which is known to be an important confounder in studies of antidepressants (Walsh et al. 2002). Therefore, D01 is given less weight.

The pivotal randomized controlled trial of 10 weeks (standard trial length for efficacy determination of an antidepressant medication) failed to demonstrate statistically significantly superior outcomes greater than sham treatment (15% versus 10%, p = 0.31 (Fisher's exact)), in contrast to the 30.5% reported response rate of the 10 week antecedent study D01. D02 was blinded, though adequacy of blinding was not verified in spite of patients sometimes being able to somatically detect active VNS treatment. The somatic detection of stimulus could encourage an expectation of response in patients who are optimistic about a new treatment (FDA Clinical Memorandum). The D02 observational study (the continuation of the D02 randomized trial) did not include the 21 sham treated (placebo) patients whose HRSD scores improved so much that they did not meet the criteria to continue in the long-term phase, thus illustrating either the natural course of the disease, where symptoms wax and wane and there may be spontaneous remission, or the placebo effect. D02 results from the observational study have been reported for 12 months, with increasing response rates. Walsh et al noted in their review of placebo response in studies of major depression, "The length of randomized controlled trials has increased, and we found, as have others, that the proportion of patients responding to placebo increases with trial length. Presumably, this association reflects both the cumulative effects of the nonspecific interventions inherent in clinical trials and a longer period during which spontaneous recovery could occur" (Walsh et al. 2002). Khan et al. in their examination of FDA data from randomized controlled trials concluded, "First, it strongly suggests that placebo-controlled trials are critical for evaluating the efficacy of treatment in this area. If clinical trial design manipulations can change symptom reductions from less than 27% in one trial to more than 61% in another, then certainly no absolute numerical cutoff will suffice for a determination of efficacy" (Khan et al. 2003). Less weight is given to studies in the area of depression that are not designed to guard against the placebo effect.

Other issues add to the difficulties with interpreting the D02 studies: inconsistent reporting of data between publications and FDA public documents; lack of rigor in patient selection; measures and endpoints that are clinically ambiguous; and concomitant adjustment of other treatments (D02 long-term study). Detailed inclusion/exclusion criteria were listed only in the FDA Clinical Memorandum. It is unclear why certain criteria were chosen, and some criteria appeared subject to broad interpretation. Inclusion criteria included: "…must have had an unsatisfactory response to at least two adequate trials of different classes of antidepressant medication, but not more than six, regardless of antidepressant category based on participant/family interviews, medical records, and, when available, pharmacy records" (Rush et al. 2005a). The attainment of antidepressant therapy history is questioned: in a recent study, the accuracy of patients recalling prior treatment with antidepressants revealed that about 80% remembered monotherapy correctly, while only 25% recalled augmentation therapy correctly (Posternak et al. 2003); a medical record review of Medicare patients receiving mental health services revealed many medical records were found to lack adequate documentation, with no documentation for billed visits in some cases (Office of Inspector General 2001). In D01, other mood disorder treatments included "phototherapy and other types of alternative treatment (e.g., St. John's wort, flaxseed oil, and fish oil)." It is not clear if the meaning of "regardless of antidepressant category" is similar between D01 and D02. It is not clear why the number of treatments was capped at six for inclusion. Patients with clinically significant suicide risk were excluded. Both criteria raise the question of what is intended by the study definition of treatment resistant depression. Using the various symptom scales for a total score to represent true patient benefit is problematic. The assumptions that these rating scales are based on have not been verified (Faravelli 2004). Response, remission, recovery, relapse and recurrence do not have standardized, empirical definitions and are subject to arbitrary interpretation. A recent article suggested, "The Task Force recommended that response criteria be met for 3 consecutive weeks to take into account error in the assessment of symptomatology and unstable symptomatic fluctuations. Requiring that response criteria be met for a reasonable period of time guards against miscategorizing transient improvement as a clinically significant benefit (i.e., a response)" (Rush et al. 2006). Significant confounding was introduced to the examination of results for the variable of interest when there was concurrent optimization of other treatments that may vary from site to site, clinician to clinician; with this random approach, one cannot be confident of which treatment, or combination of treatments, caused the clinical change.

After the results of the randomized clinical trial were known, the D02/D04 comparison study was conceived (BCBS). In addition to the methodologic issues of the D02 long-term phase mentioned above, these issues further limit the conclusions that can be drawn from any comparison of these two studies:

- Although both D02 and D04 were available to enroll subjects at similar time periods, almost all D04 patients entered after D02 was closed for enrollment (FDA Clinical Memorandum). Only 10 D04 patients enrolled while D02 was open. The FDA Clinical Memorandum stated that patient expectation for participating in an investigational study for a new therapy may have been greater than for participation in the standard treatment study. Enhanced expectation in patients hoping to enter into treatment for a new therapy could lead to improved response in D02 as compared to D04, thus leading to a potential for selection bias.
- Uncertainty in poolability of results. Twelve of 22 sites participated in both D02 and D04, one site in D04 only. Usual care may have been very different based on site, as it is known that the treatment of depression is variable. As an example, in a study by Niklson et al, treatment center did make a difference in outcome (Niklson, Reimitz 1997).
- Inclusion and exclusion criteria are not the same, thus leading to a potential for selection bias.
- There was an imbalance between groups in the 17 measured baseline variables. More D02 patients had received ECT during their lifetimes, and more patients who received ECT during the current MDE. There were more patients in the D04 population with greater than 10 lifetime episodes of depression. This again introduces the potential for selection bias.
- A propensity score used 17 baseline variables attempting to balance baseline characteristics. Propensity scores can only adjust for observed covariates, unmeasured variables can not be accounted for. The groups may not be comparable on important unmeasured variables.

- The only measure (of five) to suggest a benefit in the D02 acute-phase trial (IDS-SR – a secondary measure) was chosen for the primary measure for the D02/D04 comparison, after results from the acute-phase trial were available. Choosing an outcome in this fashion can increase the risk for false positive results. The IDS-SR is a far less frequently used measure than the HRSD. The FDA statistical reviewer did not find good concordance between the HRSD-24 and the IDS-SR (FDA Statistical Summary Review).
- The authors used IDS-SR data in a repeated measure analysis attempting to compare the rate of improvement over time between the two treatments, instead of comparing the proportion of subjects that met their previous response criteria. It is unclear if this value represents real clinical improvement for the patient, or if the measured difference between the two slopes represents a true clinical difference for an individual patient. When surrogate outcomes of uncertain clinical significance are chosen by investigators in the design of a study, those outcomes are rarely useful in making a determination of reasonable and necessary.
- As the FDA statistical reviewer concluded, “No minimum clinically detectable difference in two slopes or mean HRSD-24 or IDS-SR was defined at the study design stage in order to estimate the required sample size with pre-specified power, type I error, estimated variability of the data, number of follow-up visits, and correlation among repeated measures.” Post-hoc consideration of these design issues increases the risk for false positive results.
- Changes in antidepressant treatment were allowed in D02 and D04, so it is difficult to definitively attribute improvement to VNS. When an attempt was made using a censoring analysis employing a LOCF approach, the FDA did not find a statistically significant difference between D02 and D04, though the sponsor found otherwise. Unclear, inconsistent findings cloud interpretation of patient treatment results.
- No reason is given why adverse medical events were not reported in D04, but as a direct result any reported adverse events from D02–the safety of this therapy–can not be compared between the two groups.

Furthermore, the FDA statistical summary review stated, “Due to above statistical issues, it is unclear whether the effectiveness claim of D-02 over D-04 group has been demonstrated.”

In summary, statistical manipulation does not compensate for a poorly designed study. Upon examination, the comparison of these two observational trials provides little evidence that a patient will experience a health benefit as a direct result of VNS therapy.

Demographics of patients in the D02 trial revealed a mean age of 46.3 years, with 97% of the participants listed as Caucasian. Comorbidities that can be commonly associated with depression, such as axis I (other than mood disorders), axis II comorbidities and general medical comorbidities, were not reported. It is unclear how results from trials of patients without reported significant comorbidities can be generalized to many clinical populations, including older adults in Medicare.

In a recent study by Perlis et al of industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry, they found; “Industry sponsorship and author conflict of interest are prevalent and do appear to affect study outcomes,” and, “Given this prevalence and the potential influence on the general psychiatric literature, it will be critical to obtain a better understanding of the ways in which industry funding or the presence of conflict of interest influences the design, conduct, and/or reporting of clinical trials” (Perlis et al. 2005).

Based on the current evidence, CMS does not believe there is a treatment effect directly attributable to VNS therapy for TRD.

Conclusions

Depression is a disorder whose etiology is not completely understood. Comorbidities are common with this disorder and can complicate treatment. Treatments of proven effectiveness do exist, though there is little information to guide clinicians in medication selection depending on individual patient characteristics, and what to do if initial treatment fails. Practical clinical trials in psychiatry are now beginning to occur. “Although much progress has been made, clinical practice in psychiatry remains far from being evidence based, with heterogeneity in practice predicting heterogeneity in quality of care” (March et al. 2005). The Surgeon General’s report on mental health (1999) emphasizes a vision for the future: “high-quality research is a potent weapon against stigma, one that forces skeptics to let go of misconceptions and stereotypes concerning mental illness and the burdens experienced by persons who have these disorders.”

The concept of treatment resistant depression is vaguely defined, subject to varying determination, and until a scientifically valid definition exists is of little help in treatment selection for individual patients. In the evaluation of antidepressant treatments, randomized, controlled trial design is particularly important due to the natural course of the disease, where symptoms wax and wane and there may be spontaneous remission, and the placebo effect. The pivotal randomized controlled trial of VNS, subsequent to a pilot study, failed. Further observational study and statistical manipulations of data from this group of patients provides little weight in the determination of reasonable and necessary. CMS does not believe there is a treatment effect directly attributable to VNS therapy based on the current evidence.

**IX. Proposed Decision**

CMS is proposing that there is sufficient evidence to conclude that vagus nerve stimulation is not reasonable and necessary for treatment of resistant depression. Accordingly, we propose to issue a national noncoverage determination for this indication.

We are requesting public comments on this proposed determination pursuant to Section 731 of the Medicare Modernization Act. We are particularly interested in comments that include new evidence we have not reviewed here. After considering the public comments and any additional evidence we will make a final determination and issue a final decision memorandum.

**Appendices** [PDF, 278KB]

---

[1] increased or decreased bodily movement triggered by mental activity.

[2] Two mechanisms have been identified for the current antidepressant medications: inhibition of serotonin or norepinephrine reuptake transporters; and, inhibition of monoamine oxidase by monoamine oxidase inhibitors (Nestler et al. 2002). Interestingly, “several generations of research have failed to provide convincing evidence that depression is caused by abnormalities in the brain’s serotonin or neuroepinephrine systems” (Nestler et al. 2002).

Back to Top

# Bibliography

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC, American Psychiatric Association, 1994.

American Psychiatric Association, Practice Guideline for treatment of patients with major depressive disorder [internet]. 2nd edition, April 2000. Available at: http://www.psych.org/psych_pract/treatg/pg/MDD2e_05-15-06.pdf

Bagby RM, Ryder AG, Schuller DR et al. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry 2004; 161 (12):2163-77.

Baghai TC, Moller HJ, Rupprecht R. Recent progress in pharmacological and non-pharmacological treatment options of major depression. Curr Pharm Des 2006; 12 (4):503-15.

Bech P. Rating scales in depression: limitations and pitfalls. Dialogues Clin Neurosci 2006; 8 (2):207-15.

Benedetti F, Mayberg HS, Wager TD et al. Neurobiological mechanisms of the placebo effect. J Neurosci 2005; 25 (45):10390-402.

Blue Cross Blue Shield Association Technology Assessment. Vagus Nerve Stimulation for treatment-resistant depression. [internet] 2005. Available at: http://www.bcbs.com/tec/vol20/20_08.html

Blue Cross Blue Shield Association Technology Assessment. Vagus Nerve Stimulation for treatment-resistant depression. [internet] 2006. Available at: http://www.bcbs.com/tec/vol21/21_07.html

Cadieux RJ. Practical management of treatment-resistant depression. Am Fam Physician 1998; 58 (9):2059-62.

California Technology Assessment. Vagus nerve stimulation for treatment-resistant depression. [internet] 2006. Available at: http://www.ctaf.org/content/assessments_pdf/65198186099_VNS%200206%20W.pdf;jsessionid=FC0938AF5E7A058C9DB7EA1C7473A81D

Carmody TJ, Rush AJ, Bernstein I et al. The Montgomery Asberg and the Hamilton ratings of depression: A comparison of measures. Eur Neuropsychopharmacol 2006; 16 (8):601-11.

Depression: Management of depression in primary and secondary care, National Clinical Practice Guideline Number 23, National Collaborating Centre for Mental Health commissioned by the National Institute for Clinical Excellence. 2004. Available at: http://www.nice.org.uk/page.aspx?o=236667

Depression in Primary Care: detection and diagnosis Volume 1. Detection and diagnosis, clinical Guideline number 5, AHCPR Publication No. 93-0550: April 1993. Agency for Healthcare Research and Quality, Rockville, MD. Available at: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat6.chapter.1448

Dunner DL, Rush AJ, Russell JM et al. Prospective, long-term, multicenter study of the naturalistic outcomes of patients with treatment-resistant depression. J Clin Psychiatry 2006; 67 (5):688-95.

Elger G, Hoppe C, Falkai P et al. Vagus nerve stimulation is associated with mood improvements in epilepsy patients. Epilepsy Res 2000; 42 (2-3):203-10.

Faravelli C. Assessment of psychopathology. Psychother Psychosom 2004; 73 (3):139-41.

Fava M. Diagnosis and definition of treatment-resistant depression. Biol Psychiatry 2003; 53 (8):649-59.

Fleck MP, Horwath E. Pharmacologic management of difficult-to-treat depression in clinical practice. Psychiatr Serv 2005; 56 (8):1005-11.

Fochtmann LJ, Gelenberg AJ: Guideline Watch: Practice Guideline for the Treatment of patients with major depressive disorder, 2nd edition. Arlington, VA: American Psychiatric Association, 2005. Available online: http://www.psych.org/psych_pract/treatg/pg/MDD.watch.pdf

George MS, Rush AJ, Marangell LB et al. A one-year comparison of vagus nerve stimulation with treatment as usual for treatment-resistant depression. Biol Psychiatry 2005; 58 (5):364-73.

Gildenberg PL. The birth of stereotactic surgery: a personal retrospective. Neurosurgery 2004; 54 (1):199-207; discussion -8.

Groves DA, Brown VJ. Vagal nerve stimulation: a review of its applications and potential mechanisms that mediate its clinical effects. Neurosci Biobehav Rev 2005; 29 (3):493-500.

Hansen RA, Gartlehner G, Lohr KN et al. Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. Ann Intern Med 2005; 143 (6):415-26.

Hasin DS, Goodwin RD, Stinson FS et al. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 2005; 62 (10):1097-106.

*Improving Quality of Care for People With Depression*. Translating Research Into Practice. Fact Sheet. AHRQ Publication No. 00-P020, January 2000. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/research/deprqoc.htm

Keller MB. Issues in treatment-resistant depression. J Clin Psychiatry 2005; 66 Suppl 8:5-12.

Kessler RC, Berglund P, Demler O et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). Jama 2003; 289 (23):3095-105.

Khan A, Detke M, Khan SR et al. Placebo response and antidepressant clinical trial outcome. J Nerv Ment Dis 2003; 191 (4):211-8.

Khan A, Khan S, Brown WA. Are placebo controls necessary to test new antidepressants and anxiolytics? Int J Neuropsychopharmacol 2002; 5 (3):193-7.

Kupfer DJ, Frank E. Placebo in clinical trials for depression: complexity and necessity. Jama 2002; 287 (14):1853-4.

Licht RW, Qvitzau S, Allerup P et al. Validation of the Bech-Rafaelsen Melancholia Scale and the Hamilton Depression Scale in patients with major depression; is the total score a valid measure of illness severity? Acta Psychiatr Scand 2005; 111 (2):144-9.

Marangell LB, Rush AJ, George MS et al. Vagus nerve stimulation (VNS) for major depressive episodes: one year outcomes. Biol Psychiatry 2002; 51 (4):280-7.

March JS, Silva SG, Compton S et al. The case for practical clinical trials in psychiatry. Am J Psychiatry 2005; 162 (5):836-46.

Matthews K, Eljamel MS. Vagus nerve stimulation and refractory depression: please can you switch me on doctor? Br J Psychiatry 2003; 183:181-3.

Mayberg HS, Silva JA, Brannan SK et al. The functional neuroanatomy of the placebo effect. Am J Psychiatry 2002; 159 (5):728-37.

Medicare Part B Payments for Mental Health Services, Department of Health and Human Services, Office of Inspector General, OEI-03-99-00130, May 2001. Available at: http://oig.hhs.gov/oei/reports/oei-03-99-00130.pdf

Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979; 134:382-9.

Nahas Z, Marangell LB, Husain MM et al. Two-year outcome of vagus nerve stimulation (VNS) for treatment of major depressive episodes. J Clin Psychiatry 2005; 66 (9):1097-104.

Nelson JC, Portera L, Leon AC. Assessment of outcome in depression. J Psychopharmacol 2006; 20 (4 Suppl):47-53.

Nemeroff CB. Augmentation strategies in patients with refractory depression. Depress Anxiety 1996; 4 (4):169-81.

Nestler EJ, Barrot M, DiLeone RJ et al. Neurobiology of depression. Neuron 2002; 34 (1):13-25.

Oquendo MA, Kamali M, Ellis SP et al. Adequacy of antidepressant treatment after discharge and the occurrence of suicidal acts in major depression: a prospective study. Am J Psychiatry 2002; 159 (10):1746-51.

Perlis RH, Perlis CS, Wu Y et al. Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. Am J Psychiatry 2005; 162 (10):1957-60.

Posternak MA, Zimmerman M. How accurate are patients in reporting their antidepressant treatment history? J Affect Disord 2003; 75 (2):115-24.

Prudic J, Olfson M, Marcus SC et al. Effectiveness of electroconvulsive therapy in community settings. Biol Psychiatry 2004; 55 (3):301-12.

Rehm LP, O'Hara MW. Item characteristics of the Hamilton Rating Scale for Depression. J Psychiatr Res 1985; 19 (1):31-41.

Rush AJ, George MS, Sackeim HA et al. Vagus nerve stimulation (VNS) for treatment-resistant depressions: a multicenter study. Biol Psychiatry 2000; 47 (4):276-86.

Rush AJ, Kraemer HC, Sackeim HA et al. Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology 2006; 31 (9):1841-53.

Rush AJ, Marangell LB, Sackeim HA et al. Vagus nerve stimulation for treatment-resistant depression: a randomized, controlled acute phase trial. Biol Psychiatry 2005a; 58 (5):347-54.

Rush AJ, Sackeim HA, Marangell LB et al. Effects of 12 months of vagus nerve stimulation in treatment-resistant depression: a naturalistic study. Biol Psychiatry 2005b; 58 (5):355-63.

Rush AJ, Thase ME, Dube S. Research issues in the study of difficult-to-treat depression. Biol Psychiatry 2003; 53 (8):743-53.

Sackeim HA, Rush AJ, George MS et al. Vagus nerve stimulation (VNS) for treatment-resistant depression: efficacy, side effects, and predictors of outcome. Neuropsychopharmacology 2001; 25 (5):713-28.

Salzman C. Treatment-refractory depression. JAMA 2006;296:167-8.

Schatzberg AF, Kraemer HC. Use of placebo control groups in evaluating efficacy of treatment of unipolar major depression. Biol Psychiatry 2000; 47 (8):736-44.

Stimpson N, Agrawal N, Lewis G. Randomized controlled trials investigating pharmacological and psychological interventions for treatment-refractory depression. Systematic review. Br J Psychiatry 2002; 181:284-94.

Taylor WD, Doraiswamy PM. A systematic review of antidepressant placebo-controlled trials for geriatric depression: limitations of current data and directions for the future. Neuropsychopharmacology 2004; 29 (12):2285-99.

Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. J Clin Psychiatry 1997; 58 Suppl 13:23-9.

Thase ME, Treatment of severe depression. J Clin Psychiatry 2000;61[suppl1]:17-25.

Treatment of major depression Volume 2. Treatment of major depression, clinical practice Guideline number 5, AHCPR Publication No. 93-0551: April 1993. Agency for Healthcare Research and Quality, Rockville, MD. Available at: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat6.chapter.15593

U. S. Food and Drug Administration Center for Devices and Radiological Health. [internet]. 2004 Final Statistical Summary Review for PMA P970003/S50 (Original and Various Amendments), Vagus Nerve Stimulator (VNS) Therapy System for Depression, Cyberonics, Inc. Available at: http://www.fda.gov/OHRMS/DOCKETS/AC/04/briefing/4047b1_00_a_FDA%20Statistical%20Review%20Memo.pdf

U. S. Food and Drug Administration Center for Devices and Radiological Health. [internet]. 2005. Summary of Safety and Effectiveness Data. Available at: http://www.fda.gov/cdrh/PDF/p970003s050b.pdf

U. S. Food and Drug Administration Center for Devices and Radiological Health. [internet]. 2004. Clinical Review Memorandum..

Available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/4047b1_00_FDA%20Clinical%20Review%20Memo.doc.

U.S. Department of Health and Human Services. Mental Health: A Report of the Surgeon General. [Internet] Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes for Health, National Institute of Mental Health, 1999. Available from: http://www.surgeongeneral.gov/library/mentalhealth/home.html

Vagus Nerve Stimulation, Aetna Clinical Policy Bulletins August, 2006. Available at: http://www.aetna.com/cpb/data/CPBA0191.html

Veterans Administration Technology Assessment Program, Office of Patient Care Services. Outcome measurement in major depression [internet]. 2004. Boston, MA. Available at:

http://www.va.gov/vatap/pubs/Depressionfinal3-05.pdf

Walsh BT, Seidman SN, Sysko R et al. Placebo response in studies of major depression: variable, substantial, and growing. JAMA 2002; 287 (14):1840-7.

Wang PS, Schneeweiss S, Brookhart MA et al. Suboptimal antidepressant use in the elderly. J Clin Psychopharmacol 2005; 25 (2):118-26.

Zajecka JM. Treating depression to remission. J Clin Psychiatry 2003; 64 Suppl 15:7-12.

Back to Top